

UNIVERSIDADE DE LISBOA
FACULDADE DE LETRAS
DEPARTAMENTO DE LINGUÍSTICA GERAL E ROMÂNICA



CLAUSE STRUCTURE IN SANTOME

TJERK HAGEMEIJER

DOUTORAMENTO EM LETRAS
(Linguística Geral)

2007

Table of Contents

<i>Acknowledgements</i>	<i>i</i>
<i>Abbreviations</i>	<i>iv</i>
<i>List of Tables and Map</i>	<i>vi</i>
<i>Abstract</i>	<i>vii</i>
<i>Resumo</i>	<i>ix</i>
1. INTRODUCTION	1
1.1 Origins	1
1.2. Language name and orthography	2
1.3. Corpus	2
1.4. Theoretical framework	4
1.5. Outline of the dissertation	4
2. SUBJECT PRONOUNS AND THE SUBJECT POSITION	9
2.1. Introduction	9
2.2. Subject pronouns in previous works	9
2.2.1. Late 19 th century descriptions	10
2.2.2. Valkhoff (1966)	11
2.2.3. Ferraz (1979)	12
2.2.4. Schang (2000)	13
2.3. A new account of subject pronouns	14
2.3.1. Respectful pronouns	15
2.3.2. Unmarked/anaphoric <i>a</i>	17
2.3.3. Special uses of third singular <i>ê</i>	18
2.3.4. Expletive <i>kwa</i>	20
2.4. Evidence for a strong/weak bias	21
2.4.1. DP coordination and disjunction	21
2.4.2. Isolation	22
2.4.3. Focus	23
2.4.4. Apposition	25
2.4.5. Wh-in-situ	25
2.4.6. Subject doubling	25
2.4.7. Exclusiveness constructions	26
2.4.8. Modification	26
2.4.9. Summary	26
2.5. Phonologically reduced pronouns	27
2.5.1. Second person singular <i>ô</i>	27
2.5.2. First person singular <i>am</i> and third person plural <i>nen</i>	29
2.6. Weak pronouns: syntax or phonology?	32
2.6.1. Subject pronouns	32
2.6.2. Object pronouns	38
2.6.3. Pseudo-reflexive pronouns	40
2.6.4. Looking for a host	43

2.7. Strong pronouns as topics	46
2.8. Pronouns at the syntax-phonology interface	49
2.8.1. The internal structure of pronouns	51
2.8.2. Summary	54
2.9. Santome and the null subject parameter	55
2.9.1. The theory of <i>pro</i> -drop	55
2.9.2. Against <i>pro</i> -drop in Santome	56
2.9.2.1 Existential constructions	57
2.9.2.2. Unaccusative <i>xiga</i> ‘to arrive’	59
2.9.2.3. Copular verb <i>fika</i>	61
2.9.2.4. A raising verb?	62
2.9.2.5. Discourse-bound topic drop	63
2.9.2.6. Enumeration & coordination	65
2.9.2.7. Serial verbs	68
2.9.3. Interpreting the <i>pro</i> -drop data	69
2.9.4. Summary	71
3. THE EXTENDED VERB PHRASE	73
3.1. Introduction	73
3.2. The verb phrase and verb movement	73
3.2.1. Verb movement	74
3.2.1.1. Adverb placement	75
3.2.1.2. Quantifier float	77
3.2.1.3. Double object constructions	79
3.2.1.4. Allomorphic variation: <i>ba</i> and <i>be</i> ‘to go’	83
3.2.1.4.1. Case-marking	85
3.2.1.4.2. The comitative-goal shift	90
3.2.1.4.2.1. Scrambling and verb movement	93
3.2.1.4.2.2. DP-adjunction	95
3.2.1.4.2.3. Right-adjunction to VP	96
3.2.1.4.3. <i>Ba</i> and <i>be</i> and the lexicon	97
3.2.2. Summary	99
3.3. The TMA-system	99
3.3.1. Earlier proposals on tense and aspect in Santome	100
3.3.2. Tense and aspect markers	103
3.3.2.1. <i>Ka</i>	104
3.3.2.2. <i>Sa ka ~ ska ~ xka</i>	106
3.3.2.3. Non-marked verb	108
3.3.2.4. <i>Tava ~ ta</i>	109
3.3.2.5. <i>Tava ka ~ ta ka</i>	109
3.3.2.6. <i>Ka sa ka~ka ska</i>	110
3.3.2.7. <i>Ka ka</i>	111
3.3.2.8. <i>Sa ska</i>	113
3.3.2.9. <i>Tava ka ska, tava ka ka, tava sa ska</i>	116
3.3.2.10. Summary	117
3.3.3. Stativity	118
3.3.4. Relative tense	129
3.3.5. Projecting aspect	131
3.3.6. Projecting tense	144
3.3.7. Summary	161
3.3.8. Mood and modality	164
3.3.8.1. Mood in the TMA domain	164
3.3.8.2. High modality	166
3.3.8.2.1. <i>Sela</i>	166
3.3.8.2.2. <i>Milhon</i>	169
3.3.8.3. Summary	171

4. NEGATION IN SANTOME	173
4.1. Introduction	173
4.2. Sentence negation: the data	173
4.2.1. Contexts with Neg1...Neg2	174
4.2.1.1. Simplex sentences	174
4.2.1.2. Complex sentences	178
4.2.1.2.1. Embedding with Neg2 in sentence-final position	178
4.2.1.2.2. Embedding with Neg2 in clause-final position	182
4.2.2. Contexts without Neg2	185
4.2.3. Contexts without Neg1	189
4.2.4. Summary	190
4.3. The status of the negation markers in Santome	191
4.3.1. The status of Neg1	191
4.3.2. The status of Neg2	197
4.3.3. The relation between <i>na</i> and <i>fa</i>	198
4.3.4. Summary	201
4.4. Previous analyses of discontinuous negation	201
4.4.1. Zanuttini (1994, 1997)	202
4.4.2. Haegeman (2002b)	203
4.4.3.1. A short note on NEG 1 in Afrikaans	210
4.4.4. Aboh (2004, forthc.)	214
4.4.5. Summary	218
4.5. Towards an analysis	218
4.5.1. Adjunction and peripherality	219
4.5.1.1. The distribution of adverbials	220
4.5.1.1.1. Focus adverbs	220
4.5.1.1.2. Time-related adjuncts	228
4.5.1.1.3. Durational adjuncts	232
4.5.1.2. The distribution of causal adjuncts	235
4.5.1.3. Coordination	240
4.5.2. Summary	244
APPENDIX TO CHAPTER 4	247
1. Introduction	247
2. Negative Concord	247
2.1. N-words	247
2.2. Minimizers	253
3. Negative polarity and underspecification	255
3.1. Nê ‘nor, and, (not) even’	255
3.2. Antawo ‘(not) yet’	257
4. Diachronic and comparative perspectives	258
4.1. 19 th century sources	258
4.2. Negation and emphasis	263
4.3. Negation in the Gulf of Guinea creoles	268
4.3.1. Lung’ie	268
4.3.2. Fa d’Ambô	269
4.3.3. Ngola	271
4.4. Reconstructing negation in the GGC	272
4.5. Typology and transfer	274
4.5.1. Portuguese	274
4.5.2. Edoid	276
4.5.3. Western Bantu (area H)	277
5. Summary	283

5. CONCLUSIONS	285
5.1. Synthesis of the findings	285
5.2. Issues for further research	290
References	295

Acknowledgements

The preparation of this dissertation began officially in 2001 but it is better seen as the result of a decade of my linguistics studies and research on Santome. This decade has surely many more chapters and sections than this dissertation, among which advisors, informants, colleagues, friends and family. Fortunately, these professional and social ‘guilds’ do not have rigid boundaries and all of a sudden advisors and informants become friends, friends become advisors or informants, and so on. But let me attempt to introduce some order in this complex network of relations.

Above all, I’m greatly indebted to my advisor Inês Duarte. From the very beginning, she has shown me the way in syntax and stimulated my research on Santome. Her admirable qualities as a professor and motivational manager always kept me going in search of new and in-depth answers. At the end of our meetings, after thanking her for her excellent ideas, proposals and support, she would always say: “don’t thank me, it’s my job”. Therefore I will now take advantage of this opportunity to undo all her ‘unthankyous’ and express my deep appreciation for providing me a unique learning experience for so many years now.

I met my co-advisor Tonjes Veenstra in Los Angeles at my first conference abroad. Without knowing each other, he entered the conference room, sat next to me and asked me: “Are you Tjerk?” Ever since, we have had the opportunity to discuss the syntax of creole languages at live or long-distance occasions. I’m very grateful for his deep interest in my work, for his constructive criticisms and, on a more practical side, for sending me so many papers which I would otherwise have trouble accessing.

Between 2001 and 2005, my research was funded by Fundação para a Ciência e Tecnologia. Their grant - BD/3159/2000 - allowed me to discuss my work in a number of national and international meetings and to carry out fieldwork in S. Tomé.

Conducting fieldwork on a lesser known language has the benefit that you never get the feeling that you are stuck with a pile of paperwork and a computer. The fieldwork experience on S. Tomé was much more than just a number of academic trips in search

of data. First, I want to express my deep gratitude to Caustrino Alcântara, who played a crucial role in collecting new data with informants, transcribing and solving many of my linguistic questions. His heart for the matter and language knowledge and intuitions made things much easier for me, the “muçulmano”. I also thank Pascoal de Sousa for being such a great and efficient help with the transcriptions.

The input for these transcriptions during my stays on the island in 2001 and 2002 was provided by many people (often talking at the same time, to make things even harder for me!): the people from Ototo, especially Ernestino Espírito da Silva, the people of Mateus Angolar, sun Lomba and san Patlixa from Madre de Deus, sun Katxi from Bobo Forro and especially the late Serafina Bonfim, daughter of the legendary Francisco Bonfim, better known as Faxiku Bêbêzawa, for her wonderful story-telling and for letting me copy the controversial pamphlets written in creole by her father. After the unique experience I had in tape-recording “Mese” Sabino during my first trip in 1997, I found him in a much debilitated state in 2001. Nothing is more true than the words he said about tape-recording him: *Sabino ka pô môlê, a ka lembla Sabino* (‘I may die, but I will be remembered’). To me, this is almost a prophecy.

Especially, but not only, the following people made my day-to-day life in S. Tomé an even more enjoyable experience: Bia, Caustrino, Gastão, Leonel, Luís & Bibi, Pascoal and Rodorico.

At home in Lisbon, I collected substantial amounts of elicited data from Jerónimo Pontes, whose great intuitions and critical spirit often made me revise my assumptions about the data I was seeking. I also thank Ivo Jordão for sharing his profound language knowledge with me. Ever since I began work on Santome, I have had the support of my friend Beatriz Afonso, who is one of those persons that, like me, would really like to see a better future for the creole. In S. Tomé or in Lisbon, she was always available for my queries.

As a participant in the yahoogroup for S. Tomé, I stumbled across two persons who took deep interest in my project and have been helping me by email for the last two years with my tons of questions. I’m very grateful to Alcídio Pereira for critical answers to my queries and a very, very special thank you goes to Conceição Lima, who helped me out with the data far more than one can ever expect from a person, especially in the final stages of this work.

Many, many thanks to all the people from S. Tomé who I have worked with over the past decade. They make me feel deeply committed to Santome and I sincerely hope I can continue to give back what has been so generously given to me.

Philippe Maurer has been the best business partner one could possibly wish for. Thanks to his generosity, my knowledge of the Gulf of Guinea creoles has grown immensely and my bookshelves gained quite a few exotic titles. But above all, he is a dear friend, the kind of person one always loves to have around.

I owe a special thanks to Alan Baxter, Tom Güldemann and Armin Schwegler, with whom I have been in close contact during the preparation of my dissertation. I have greatly benefited from their linguistic experience, friendship and help in many different respects. I would further like to thank Enoch Aboh, Arthur Bell, Hans den Besten, Arlindo Caldeira, Ben Elugbe, Philip Havik, Gerardo Lorenzino, Ota Ogie, Mikael Parkvall, Marike Post, Nicolas Quint, Jorge Rocha, Ron Schaefer, Scott Schwenter, Gerhard Seibert, Norval Smith and Maria Trovoada for broadening my learning experience in many diverse ways.

To the knights of the rectangular table of room 31, my dear friends Nélia Alexandre, Ana Lúcia Santos and Nuno Soares, and close colleagues Madalena Colaço, Ana Luisa Costa, Fernanda Gonçalves, Matilde Miguel and Gaétan de Saint-Moulin, and to Fernanda Pratas, my sincere appreciation for the friendship and all the pleasant yet very constructive debates. I guess you all know by now how to kill a pig in Santome.

Friends will be friends: Tom, Jelmer and Inês and especially Zé and Ana for always taking such good care of me.

During the preparation of this dissertation, my dear grandmothers and grandfather left for a better place and are still missed. Fortunately they left behind the best parents in the world. Pa & Ma, Hugo, Sabine and little Hugo, I can't thank you for always being there; your being here would be much better.

Lena and Laura, you're the magic in my life, my home, my haven.

Abbreviations

Adj(P)	Adjective (Phrase)
Adv(P)	Adverbial (Phrase)
AFR	Afrikaans
Agr(P)	Agreement (Phrase)
Agro(P)	Object Agreement (Phrase)
ANT	Anterior
ASP	Aspect
Asp(P)	Aspect (Phrase)
Aux(P)	Auxiliary (Phrase)
COMP	Complementizer
Conj(P)	Conjunction (Phrase)
C(P)	Complementizer (Phrase)
DEF	Definite
DEM	Demonstrative
DO	Direct Object
DOC	Double Object Construction
D(P)	Determiner (Phrase)
EMPH	Emphasis
EPP	Extended Projection Principle
EXCL	Exclamative
EXPL	Expletive
FA	Fa d' Ambô (Annobonese)
Fin(P)	Finite (Phrase)
FOC	Focus
Foc(P)	Focus (phrase)
FUT	Future
GGC	Gulf of Guinea creoles
IMP	Impersonal
IMPER	Imperative
INFL	Inflection
INT	Interrogative
IO	Indirect Object
IP	Inflection Phrase
IRR	Irrealis
LCA	Linear Correspondence Axiom
LF	Logical Form
LU	Lung'ie (Principense)
MIN	Minimizer
Mod(P)	Modal (Phrase)
NC	Negative Concord
NEG	Negation
Neg(P)	Negation (Phrase)
NG	Ngola (Angolar)
NP	Negative Polarity
NPI	Negative Polarity Item
OCC	Occurrence Feature
PDH	Parameterized Direction Hypothesis

PF	Phonological Form
PL / pl	Plural
Pol(P)	Polarity (Phrase)
POS	Possessive
PP	Preposition Phrase
PRES	Presentational Marker
PSR	Pseudo-reflexive
PST	Past
PTG	Portuguese
RED	Reduplication
REL	Relative Marker
SBJV	Subjunctive
SC	Small Clause
SCL	Subject Clitic
Spec	Specifier
SG / sg	Singular
Sp(P)	Specific (Phrase)
SP	Specific
ST	Santome (São Tomense)
SUBJ	Subject
SVC	Serial Verb Construction
TMA	Tense-Mood-Aspect
TNS	Tense
Top(P)	Topic (Phrase)
T(P)	Tense (Phrase)
VOC	Vocative
WF	West Flemish

List of Tables and Map

Chapter 2

Table 1. Subject pronouns in late 19 th century descriptions.	10
Table 2. Subject pronouns in Valkhoff (1966: 96)	11
Table 3. Disjunctive pronouns in Valkhoff (1966: 96)	12
Table 4. Subject pronouns in Ferraz (1979: 62)	12
Table 5. Subject pronouns in Schang (2000: 225)	14
Table 6. Subject pronouns in Santome	15
Table 7. Direct object pronouns in Santome	38
Table 8. Possessive and pseudo-reflexive pronouns	41
Table 9. Properties of subject pronouns	49
Table 10. Distribution of subject and object pronouns at the syntax-phonology interface	50

Chapter 3

Table 1. Complementary distribution of <i>ba</i> and <i>be</i> ‘to go’	83
Table 2. Lexical items with ‘prepositional functions’	87
Table 3. Distribution of <i>be</i> and <i>ba</i>	98
Table 4. Classification of tense and aspect markers in the GGC	101
Table 5. Value of TMA-markers according to Schang (2000: 197)	102
Table 6. Semantic functions of tense-aspect markers (with dynamic predicates)	117
Table 7. Occurrences of tense-aspect markers in the spoken corpus	118
Table 8. Functions of <i>ka</i> and \emptyset according to lexical aspect	120
Table 9. Types of stativity and tense-aspect marking	127
Table 10. Number of occurrences of the progressive in the spoken corpus	139
Table 11. Number of occurrences of the progressive in older sources	139
Table 12. Phono-syntactic properties of TMA-constructions	152
Table 13. Typology of the <i>tava</i> -constructions	157
Table 14. Number of occurrences of the past imperfective in the spoken corpus	160
Table 15. Number of occurrences of the <i>tava</i> +VP-construction in the spoken corpus	160

Appendix to Chapter 4

Table 1. Graphic representation of negation markers in Santome in the 19 th century.	261
Table 2. Negation patterns in the GGC.	273
Table 3. Preverbal negation markers in Edo and Santome.	277

Map 1: Distribution of types of negation over the Bantu language zones (Westphal 1958)	278
-------------------------------------------------------------------------------------------	-----

Abstract

This dissertation concerns the clause structure of Santome, a Portuguese-related creole language spoken on the island of São Tomé, in the Gulf of Guinea.

Chapter 2 focuses on the status of subject pronouns. I present evidence that pronouns in Santome can only be fully understood at the syntax-phonology interface. Despite the syntactic strong-weak bias of the pronominal system, I will argue that Santome does not exhibit syntactic clitics and that the cliticization phenomena should receive a phonological explanation. Complementary to this claim, I will argue that Santome is not a *pro*-drop language.

Chapter 3 is concerned with the properties of the extended VP. Standard and language-specific tests show that Santome does not exhibit verb movement, which is expected given the absence of inflectional morphology and the presence of lexicalized preverbal markers that follow the order Mood-Tense-Aspect. It can be shown that these markers have different syntactic properties that range from bound to free morphemes and reveal the fine-grained functional structure of the preverbal field. The temporal reading of clauses is obtained through the intricate relation between TP and AspP.

Chapter 4 investigates the properties of standard clausal negation, which is discontinuous in the sense that it consists of a preverbal (Neg1) and a strongly final marker (Neg2). It is argued that the preverbal marker heads a NegP that dominates the TMA-system, whereas the final marker heads a NegP inside the TMA-system. AspP-raising to [Spec, NegP2] provides the correct surface position of the final marker. The position of Neg2 also provides evidence for the classic distinction between peripheral and non-peripheral adjuncts. The latter are adjoined in a low position and pattern to the left of Neg2. Peripheral adjuncts are high adjuncts occurring to the right of Neg2, out of the scope of negation.

Keywords: Santome, subject pronouns, TMA-markers, discontinuous negation, functional projections

Resumo

Esta dissertação discute diversos aspectos da estrutura da frase em Santome, uma língua crioula falada na ilha de S. Tomé, no Golfo da Guiné, que terá começado a surgir em finais do século XV como resultado do contacto entre o português e línguas do continente africano, com particular destaque para os grupos linguísticos edoide (Nigéria) e bantoide da área H (Congo). O Santome é actualmente falado por grande parte da população são-tomense, quer como L1 quer como L2, e não tem estatuto de língua oficial.

Este estudo inscreve-se na Teoria dos Princípios e Parâmetros da Gramática Generativa (Chomsky 1955 e trabalho subsequente), com referência ocasional a desenvolvimentos posteriores neste mesmo quadro teórico (Minimalismo, Teoria das Fases) e à Abordagem Cartográfica (Rizzi 1997, Cinque 1999), e visa contribuir para um conhecimento mais aprofundado da sintaxe do Santome, visto tratar-se de um domínio largamente ignorado desde os primeiros estudos sobre esta língua, que datam de finais do século XIX. Assim, serão desenvolvidos três tópicos centrais, nomeadamente, os pronomes sujeito, o SV alargado e a negação frásica.

Nos trabalhos anteriores sobre os pronomes sujeito (Schuchardt 1888, Valkhoff 1966, Ferraz 1979, Schang 2000), a tónica recaía sobre a distinção entre pronomes fortes e fracos. Embora essencialmente correcta, esta distinção não contempla um amplo leque de factos, como será mostrado no Capítulo 2. Além de introduzir alguns pronomes anteriormente ignorados e completar a descrição de outros, a minha proposta de análise realça a importância da interface entre a fonologia e a sintaxe, consubstanciando-se numa tripartição sintáctica entre pronomes fracos, fortes e subespecificados, que, por sua vez, podem ser clíticos fonológicos ou não. Esta tipologia permite identificar novas classes pronominais para as quais há, efectivamente, evidência empírica, tais como as classes não previamente identificadas dos pronomes fortes que se comportam como clíticos fonológicos, dos pronomes fracos que não são clíticos fonológicos e clíticos fonológicos subespecificados para a sintaxe. Resulta, pois, que esta tipologia se afasta em alguns aspectos da proposta de Cardinaletti & Starke (1999).

Concluimos também que as mudanças no sistema pronominal não operam a nível do paradigma e sim a nível dos pronomes individuais. De facto, a tipologia acima esboçada pode ser redesenhada de modo a incluir um maior número de subclasses se

tivermos em conta critérios secundários, como por exemplo a direccionalidade da cliticização fonológica ou a referencialidade.

Contrariamente ao que tem sido proposto para diversas línguas crioulas desde DeGraff (1993), demonstro que o Santome não possui clíticos sintácticos, pese embora a existência de pelo menos um pronome (1sg *n*) que se aproxima significativamente dessa classificação. Como a proposta de DeGraff implica uma tipologia especial de sujeitos nulos, não há qualquer evidência de que o Santome exiba *pro-drop*, o que é corroborado pelo facto de apresentar apenas casos muito esporádicos de queda de sujeitos não argumentais.

O Capítulo 3 discute diversos aspectos relacionados com o SV alargado. Começamos por mostrar que o Santome não apresenta movimento do verbo, com base em critérios clássicos, como a posição dos advérbios e a flutuação de quantificadores, e critérios internos à língua, nomeadamente, as propriedades das construções de duplo objecto e a relação entre comitativos, pseudo-reflexivos e argumentos locativos quando ocorrem com os alomorfes *be* e *ba* ‘ir’.

No entanto, a parte mais substancial do capítulo incide sobre o sistema de T(empo)-M(odo)-(A)sp(ecto), que no casos dos crioulos do Golfo da Guiné segue a ordem MTA. Embora tenha merecido a atenção de outros autores, incluindo Bickerton (1981), as descrições existentes são lacunares e não fazem justiça à complexidade do sistema. Com base numa nova descrição, mais exhaustiva, apresentamos uma análise detalhada dos marcadores pré-verbais nucleares de tempo e aspecto, *ta~tava*, *ka*, *sa ka~ska* e \emptyset , e das respectivas combinatórias. Verifica-se, em primeiro lugar, que o Aspecto é muito mais gramatical do que o Tempo. É apresentada evidência para o facto de existirem duas projecções aspectuais (AspP) e uma temporal (TP). Além disso, existe, neste crioulo, uma correlação muito estreita entre o aspecto perfectivo (\emptyset) e o tempo passado e entre o aspecto imperfectivo (*ka*, *sa ka~ska*) e o tempo presente.

Para dar conta dos dados, propomos, por isso, um mecanismo em que T herda os traços temporais de Asp. Assumimos que a marca de tempo (*ta~tava*), que ocorre em construções [+/- anteriores], com propriedades claramente distintas, é apenas especificada para o traço [Passado]. Este valor é amalgamado com a especificação temporal herdada de Asp. Por exemplo, para uma construção [+anterior] (mais-que-perfeito) como *Zon* [_{TP} *tava* [_{AspP} \emptyset [_{VP} *kume*]]] ‘Zon tinha comido’, propomos que o traço perfectivo associado a AspP acumula como o traço [Passado] de T, resultando no valor final de mais-que-perfeito.

A parte final do Capítulo 3 mostra que o Santome possui um elemento funcional *ká* para contextos de modo. Este elemento, o núcleo de MoodP, distingue-se da marca aspectual *ka* pelo tom alto e pelo facto de anteceder a marca de tempo, motivando a ordem MTA. Será também demonstrado que existe ainda uma projecção sintáctica mais alta, presumivelmente na parte inferior da periferia esquerda, que alberga elementos modais.

Capítulo 4 investiga a negação frásica, que consiste em dois elementos descontínuos, *na...fa*, um padrão que é tipologicamente marcado. O primeiro destes elementos ocorre em posição pré-verbal, precedendo o complexo de MTA, ao passo que *fa* ocorre em final de oração. Mostramos que complementos e material de natureza XP (adverbiais) ocorrem tipicamente à esquerda de NegP2 e adjuntos oracionais tipicamente à direita. Argumentamos que os dois marcadores constituem núcleos de NegP, à semelhança do que tem sido assumido, nos últimos anos, para o Afrikaans (Bell 2004) e diversas variedades do grupo Gbe (Aboh 2004, forthc.). As propostas de análise para estas línguas defendem essencialmente que o NegP2 encabeçado pela marca final se encontra numa posição mais alta do que o NegP que ocorre primeiro na ordem linear. Nestas análises, embora em quadros teóricos diferentes, a ordem de superfície é obtida através do movimento do NegP1 e do material que selecciona para [Spec, NegP2]. Distanciamos-nos destas análises, propondo, em alternativa, que NegP2, encabeçado por *fa*, corresponde a uma projecção funcional mais baixa, situada no interior do sistema de MTA. Assim, NegP1 (*na*) c-comanda NegP2 (*fa*) e propomos que existe uma relação de Concordância a distância entre estas duas projecções. Para derivar a ordem de superfície correcta, postulamos uma operação de movimento de AspP para [Spec, NegP2]. Uma vez que complementos e adverbiais ocupam, em geral, uma posição abaixo de AspP, esta análise prediz correctamente que este tipo de material ocorrerá à esquerda de *fa* após movimento de AspP. Assumindo que há adjunção à direita, os adjuntos mais altos, tipicamente com tempo independente, ocorrem numa posição mais alta do que os dois NegPs e ocorrem, por isso, à direita de *fa*. Neste sentido, a posição de *fa* na estrutura da frase permite, crucialmente, distinguir entre adjuntos periféricos e não periféricos (Lobo 2002, 2003). Os dados de coordenação, por exemplo, mostram que os dois termos de uma coordenação baixa, por exemplo de SVs, ocorrem à esquerda de *fa*, ao passo que coordenação alta, por exemplo de NegPs, requer negação independente em cada membro coordenado.

O anexo ao Capítulo 4, por fim, apresenta dados que mostram que o Santome é uma língua de concordância negativa estrita (Zeijlstra 2004), mas que também possui diversos itens que adquirem polaridade negativa na presença da negação frásica. Na parte final do anexo, será apresentada uma perspectiva diacrónica e comparativa da negação frásica no Santome e nos outros crioulos do Golfo da Guiné. Tal como outros domínios gramaticais, os padrões de negação mostram, por um lado, vestígios do contacto linguístico entre os diversos estratos que contribuíram para a formação do Santome e, por outro, aspectos que realçam a importância da evolução interna à língua.

Palavras-chave: Santome, pronomes sujeito, marcadores de TMA, negação descontínua, projecções funcionais

1. INTRODUCTION

This dissertation is the result of in-depth research into the following topics in the clause structure of Santome: subject pronouns, the extended verb phrase and negation. These topics will be introduced in section 1.5. Before doing so, however, I will make some concise comments on issues related to this work. Section 1 succinctly introduces the origins of Santome. In section 1.2, I will provide some information on the name that I will use for the language. Section 1.3 provides an overview of the language data sources used in this dissertation. Section 1.4 briefly presents the framework used in this dissertation

1.1 Origins

Santome is a Portuguese-related creole language spoken on the island of São Tomé in the Gulf of Guinea that resulted from the contact between Portuguese and different African languages in the late 15th and early 16th century. At that time, a proto-Gulf of Guinea Creole (proto-GGC) came into being which was soon diffused into four creole varieties: **Santome** and **Ngola** spoken on the island of S: Tomé, **Lung'ie** spoken on the island of Príncipe and **Fa d'Ambô** spoken on the island of Pagalu (former Annobón). Santome, the object of this study, can be seen as the continuation in time of this proto-GGC. Linguistic, historical and genetic¹ evidence converge on the special importance of the Edoid cluster² and of Bantoid languages of area H, especially the Kongo varieties and Kimbundu, in the formative stages of a proto-Gulf of Guinea Creole (proto-GGC) (Ferraz 1979, Lorenzino 1998, Hagemeyer 1999, 2005b, Schang 2000, Hagemeyer & Parkvall 2001, Tomás et alii 2002, Hagemeyer & Güldemann 2006, Rocha et alii (under revision)). In my own work, I have argued that a founder impact of the Edoid cluster is visible in the morphosyntax of the GGC in detriment of the Bantoid cluster.

¹ Mainly because of the still incomplete genetic map of the relevant African mainland, genetics is currently unable to establish in-depth comparisons with the population of S. Tomé. However, several interesting findings have emerged. For instance, in a study on autosomal loci, namely β^s -globin haplotypes, Tomás et alii (2002) conclude that this haplotype's distribution is in agreement with the idea that the areas typically associated with the origins of the slaves, namely Bantu areas and Benin (where the latter comprises the area from Ghana to Gabon), were crucial to the peopling of the islands. Since the number of Bantu slaves imported to S. Tomé is generally considered much more significant than the number of West-Africans from the bight of Biafra/Benin (Curtin (1969), for instance, mentions a 80%-20% proportion), the average frequency of 52,3% for the Benin haplotype studied is higher than expected.

² Edo, one of the Edoid languages, was the language of the ancient kingdom of Benin that was situated in the upper Niger delta in modern Nigeria.

1.2. Language name and orthography

Santome is also known in the literature as São Tomense (also São-Tomense) or the Creole of São Tomé. Native speakers refer to their language as Santome, Lungwa Santome ('language of S. Tomé'), Lungwa Tela ('language of the country'), Diôletu ('dialect'), or Fôlô. Unlike the other native names, the latter designation also has a Portuguese counterpart, namely *Forro*, a term historically related to Portuguese (*carta de alforria* '(letter of) manumission').

To the best of my knowledge, the designation São Tomense was extensively used for the first time by Ferraz (1979) in his monograph on this language, but it should be noted that Schuchardt already refers to this language, in German, as (*Das*) Santhomensische, which is tantamount to São Tomense.

Moreover, Ferraz (1979: 8) refers to the other GGC as Principense, Annobonese and Angolar. Throughout this dissertation, I will use the most common name whereby each GGC language is known to its speakers, namely Santome, Lung'ie ('language of the island'), Fa d'Ambô ('speech of Annobón') and Ngola.

1.3. Corpus

The corpus that underlies this work is drawn from as many available sources as possible:

- transcribed tape-recordings;
- elicited data
- materials written in Santome (books, pamphlets, etc.)
- linguistic studies

It must be mentioned that Santome and the other GGC lack an official orthography. I will essentially follow the abbreviated orthography proposal of Alcântara & Hagemeyer (Ms.)³, which is very similar to the orthography used by Philippe Maurer in his work on the GGC.

³ Alcântara, C, & T. Hagemeyer (Ms.). "Proposta Ortográfica para o Santome" (presented at Colóquio Internacional sobre as Línguas Nacionais, S. Tomé, 2001).

Tape-recording and transcriptions

The tape-recording and transcriptions were made on several locations in S. Tomé in 1997 and 2001: These tape-recordings involved many people from different locations of S. Tomé, aged between 25 and 78: sun Lomba, san Patlixa (Madre de Deus); sun Katxi (Bobo Forro); Serafina Rita Bonfim, Crispim Espírito Santo (Almeirim); Caustrino Alcântara (Boa Morte); sun Sabino (Maianço); de Sousa family (Fuji Fala); Quintas family (Belém); Raimundo Santiago (Batepá); Fernando Jordão (Capela); Ernestino Espírito da Silva and friends (Ototo); sun Zon, Nankwetu, Tome, Luis Morais (Boa Entrada/Santo Amaro); the people from Mateus Angolar. Owing to the bad quality of the recordings, I was not able to transcribe Rodorico D'Alva (Monte Café), sun Raúl (Belém) and a small part of the recordings with sun Sabino. The number of transcribed words is approximately 125.000.

It should be noted that on a few occasions in Chapter 3 I use this data collection and the materials written in Santome referred to below to provide a global overview of the frequency of occurrence of TMA-markers and constructions. This use of the data is pre-statistical and only serves an informative function.

Elicited data

Primary consultants: Caustrino Alcântara, Conceição Lima, Jerónimo Pontes.

Secondary consultants: Ivo Jordão, Beatriz Afonso, Pascoal de Sousa, Alcídio Pereira.

Materials written in Santome

As mentioned, Santome lacks an official orthography, but has sporadically been used for written purposes. I will on occasion use the following written sources:

- Negreiros (1895)⁴
- Bonfim: newspaper articles from *A Liberdade* (1920ties) and pamphlets from the late 1940ties and/or early 1950ties
- Quintas da Graça (1989)
- *Revista Cultural* 1 (1990)
- Espírito Santo (1998)
- Daio (2002)

⁴ Strictly speaking, the language chapter in this work is a study, but the significant number of examples make it a unique corpus of the 19th century language.

Studies

The linguistic studies focusing specifically on the grammar of Santome are the following:

- Coelho (1880)
- Schuchardt (1888)
- Negreiros (1895)
- Valkhoff (1966)
- Ferraz (1979)
- Schang (2000)

These studies will be addressed throughout the dissertation, whenever relevant.

1.4. Theoretical framework

The research reported in this dissertation is essentially couched in the framework of the Theory of Principles and Parameters (Chomsky 1986) of the Generative Grammar tradition, with occasional references to Minimalism (Chomsky 1995) and the Cartographic Approach to language structure (Rizzi 1997, Cinque 1999). The Theory of Principles and Parameters has proved itself of great direct or indirect importance for creole studies, especially since the work of Bickerton (1975, 1981, 1984), who argues that the structure of creoles and the similarities between creole languages provide crucial insights into the human language faculty and the origins of language. Although the field of creole studies has undergone significant advances since Bickerton and has uncovered the fragilities of his seminal contribution, first and second language acquisition are still crucial in today's creole research program.

1.5. Outline of the dissertation

The topics in the clause structure of Santome investigated in this dissertation have only been superficially addressed in the small amount of previous research, and usually at an exclusively descriptive level. As a consequence, this may give the (wrong) impression that this language has a 'simple' grammar. An example of such impressions is TMA-marking, a topic that will be addressed in Chapter 3. Since the work of Valkhoff (1966),

there have been no significant improvements on this topic. Many of the TMA-constructions described in Chapter 3 have simply gone unnoticed over the decades.⁵

The lack of detailed studies of Santome should be framed within the more general picture of the history of creole studies. A few decades ago or even less the study of creole languages was still a young field, founded mostly on understudied languages that were often still trying to liberate themselves from the stigmas of the past. The true complexities and comparative differences of such languages were frequently underestimated and understated. Even an author such as Bickerton (1984: 178) was misled and suggested that most of the grammar of Saramaccan, a Surinam creole, could be explained by postulating some ten rules. Fortunately, there is also the other side of the coin. Thanks to researchers such as Bickerton, the tide keeps turning and there has been a boom in the research on creole languages. Saramaccan, Haitian or, within the ‘family’ of Portuguese-related creoles, Capeverdean, nowadays fare well in linguistics.

Santome, however, has not yet reached that privileged status and continues to be one of the lesser known creoles in the field, especially where its syntax is concerned. My primary aim is therefore to demonstrate that substantial refinements can be made with respect to the three main topics in this dissertation: subject pronouns, the extended VP and negation.

In **Chapter 2**, I will examine subject pronouns and the subject position. Descriptions of Santome have mainly focused on the syntactic distinction between weak and strong pronouns and it will be shown that this is indeed a part of the picture. However, descriptions based on syntax and a weak-strong opposition alone miss out on many important facts, since pronouns in this language are a typical case where one cannot do without the syntax-phonology interface. I will discuss whether the syntactic split between weak and strong pronouns is sufficient to account for all cases and I will propose that a phonological criterion is necessary to further refine the system.

Another goal of this chapter is to investigate whether subject pronouns in Santome are syntactic clitics, a claim that has been made for a number of creole languages and which implies that these languages constitute a special type of *pro*-drop languages.

⁵ For closely related Ngola and Lung’ie, however, in-depth descriptions of TMA-marking are available (Maurer 1995, 1997, forthc.).

Chapter 3 is a window into the functional structure of the preverbal domain. The preverbal domain of creole languages has come under the spotlight because of Bickerton's claim that creole languages exhibit TMA-systems, i.e. preverbal markers that encode (T)ense, (M)ood and (A)spect. This is arguably his last bulwark in the sense that it has best resisted his critics.

Santome and the other GGC exhibit a TMA-system but instead of TMA the markers are organized as MTA, as illustrated in the following example:

- (1) *Xi non d'ola se na ká tava ka da ku ngê*
 if 1PL of-hour SP NEG MOOD TNS ASP give with people
tamen fa, mo ngê d'oze ka vivê?
 adult NEG how people of-today ASP live
 'If we back then wouldn't have gotten along with the adults, how would today's people live?'

In the literature on creole languages, labels such as 'preverbal marker' or 'preverbal particle' have often been taken for granted. I will examine whether these elements can be subsumed under a single typology or whether their different properties motivate a differentiated analysis. Not only will it be argued that there are significant differences between, for instance, aspect markers and the tense marker, but it will be seen that even the tense marker itself exhibits different properties according to the construction in which it occurs. This raises a number of questions about the status of the projections that host these items. The functional structure of tense and aspect is also an interesting testing ground for examining how clauses obtain their temporal information. What does aspect mean for tense in a language such as Santome? Can stative and non-stative predicates be subsumed under the same type of analysis and do they exhibit the same functional structure? Are perfective and imperfective constructions birds of a feather?

Apart from the functional structure of the temporal-aspectual domain, which is the core topic of Chapter 3, the first part of the chapter discusses whether Santome exhibits verb movement. I will apply language-specific tests that *en passant* provide insights into the functional structure of the lower part of the clause, namely VP itself. Finally, at the end of the chapter I will discuss several data that are suggestive of the existence of a functional projection above TP that houses certain modal elements. .

Chapter 4 continues the research into the preverbal domain of lexical-functional projections. Santome is sometimes mentioned in the literature because of its discontinuous negations patterns, consisting of a preverbal marker *na* and a strongly final marker *fa*, which can be seen in example (1) above. It will be argued that each negation marker heads its own NegP. The NegP headed by *na*, I claim, occurs above the functional projection hosting the TMA-markers. However, my primary goal will be to examine the syntactic position of *fa*.

Assuming that *fa* heads a functional projection, there are basically two possibilities:

- (i) *fa* heads a functional projection below the NegP headed by *na*
- (ii) *fa* heads a functional projection above the NegP headed by *na*

Although double-headed negation languages have only just begun to receive attention in the generative framework, solution (ii) prevails in the literature. In the relevant analyses, the correct order is derived by moving NegP1 into the specifier of NegP2. In the case of Afrikaans, for instance, certain constructions require remnant movement. I assume that the analysis in (ii) fails to explain some important facts about the relation between the two negation markers. Any analysis of *fa* in Santome has to take into account its syntactic sensitivity to adjunct types. I will show that the distinction between peripheral and non-peripheral adjuncts is crucial to an understanding of the syntax of *fa*.

The Appendix to Chapter 4 discusses other aspects of negation in Santome, namely negative concord and cases of polarity. The final section of this appendix examines the origins and the diachronic evolution of negation in the GGC. Part of the discussion focuses on the possible relation between emphasis marker *fa~fan* and negation marker *fa*.

2. SUBJECT PRONOUNS AND THE SUBJECT POSITION

2.1. Introduction

This chapter investigates pronouns and the subject position in Santome. Earlier descriptive works on the pronominal system of Santome and the GGC in general have typically emphasized the split between two types of pronouns, labeled “bound/free” or “emphatic/non-emphatic”, *inter alia*. Careful inspection reveals, however, that (i) finer-grained distinctions are warranted, and (ii) the syntax-phonology interface plays a crucial role with respect to the workings of the pronominal system. One of the main tasks will be to assess whether the paradigm of weak pronouns in Santome behave like syntactic or phonological clitics, a topic that has been on creolists’ research agenda since the early nineties. The debate on subject clitics is intimately related to another controversial issue in creole studies: are there creoles that exhibit *pro*-drop? As will be demonstrated, both these topics have important typological implications.

The chapter is organized as follows. Section 2.2 contains a summary of previous accounts of the subject pronoun paradigm. Section 2.3 provides a new account of subject pronouns and briefly discusses several pronominal forms that have not generally been referred to in the previous literature, or the treatment of which I considered incomplete. Section 2.4 presents numerous tests in support of a split between weak and strong pronouns. Section 2.5 provides evidence for an additional class of pronouns, namely strong phonologically reduced forms. Section 2.6 investigates whether weak pronouns are best subsumed under a phonological or a syntactic approach. Section 2.7 shows that strong pronouns are topics. Section 2.8 then summarizes the properties of subject pronouns, and looks into the internal structure of weak and strong forms. Finally, section 2.9 addresses whether Santome exhibits any evidence for *pro*-drop.

2.2. Subject pronouns in previous works

In the following sections, I will briefly summarize what has been claimed about subject pronouns in previous work.

2.2.1. Late 19th century descriptions

Coelho (1880) and Schuchardt (1882) — two pioneers in creole studies — and Negreiros’ (1895) description of Santome provide important insights into the paradigm of subject pronouns. Although not fully accurate, the work of these three authors on Santome can generally be considered quite reliable when the data are compared to contemporary language data. Unlike the other two authors, Negreiros did not provide a systematic account of the pronominal paradigm. The information in Table 1 below is, therefore, a reconstitution of the pronominal paradigm based on examples that are provided throughout his language chapter. In all cases, the original orthography has been maintained.

Table 1. Subject pronouns in late 19th century descriptions.

	<i>Coelho</i>	<i>Schuchardt</i>	<i>Negreiros</i>
1sg	<i>un</i>	<i>mi, amí, amú,</i> conj. <i>um</i>	<i>n, ami</i>
2sg	<i>bô</i>	<i>bô</i>	<i>bô</i> (informal) <i>sun, san</i> (formal)
3sg	<i>ê</i>	<i>ê</i>	<i>ê</i>
1pl	<i>non</i>	<i>nom</i>	<i>non</i>
2pl	<i>nansi</i>	<i>inancé, nancé</i>	<i>inancê</i>
3pl	<i>nem</i>	<i>inen, nem</i>	<i>inêm</i>
impersonal/ anaphoric			<i>a</i>

The table shows that the authors provide identical pronominal forms. The only form that can be considered awkward in light of more recent data is Schuchardt’s *amú*, not found in other sources (contemporary or past), but perhaps he was referring to the strong reduced pronoun *am* (cf. section 2.5.2) or to 1sg *amu* in Fa d’Ambô (e.g. Post 1994). Although found in examples of Negreiros’ description, neither the formal forms for 2sg (*sun, san*) nor the impersonal/anaphoric form *a* is mentioned explicitly by any authors (on this point, see sections 2.3.1 and 2.3.2). Moreover, Schuchardt is the only author who presents variation within some forms (1sg, 2pl, 3pl) while mentioning that he was only able to find a weak/strong opposition in the 1sg form. His reference to conj. *um* is presumably an indication that he considered it a clitic. According to Schuchardt, strong

1sg only occurs after prepositions. Since this claim does not carry over to 2pl and 3pl, I assume that he considers these forms (free) variants. Somewhat surprisingly, strong 3sg *êlê* is not attested in any 19th-century description.

2.2.2. Valkhoff (1966)

It took another 70 years until the first systematized account of the subject (and object) paradigm was published. Valkhoff (1966) provides the first brief linguistic description of several aspects of Santome grammar, including an inventory of pronouns and their allomorphic and free variation. Table 2 resumes his findings in the original phonetic orthography:

Table 2. Subject pronouns in Valkhoff (1966: 96).

	<i>Singular</i>			<i>Plural</i>		
	1st	2nd	3rd	1st	2nd	3rd
Subject	<i>ami, n-</i>	<i>bo</i>	<i>e</i>	<i>nō</i>	<i>ināsé</i>	<i>iněy</i>
Allomorphs	<i>m~ɲ, mi</i>				<i>nāsé</i>	<i>ne</i>
Free variants	<i>mm~ɲɲ~nn</i>		<i>ee</i>		<i>ināse~nāse</i>	<i>ině~ine~iné</i>

Despite Valkhoff’s concern for phonological variation, there are several problems with his findings. First, I did not find any evidence for the allomorph of strong 1sg *mi*. There are contexts at word boundaries where the adjacency of a word-final *a* and *ami*’s word-initial *a* result in a contracted form, for instance *pla ami* > *pl’ami* ‘for me’. In contexts of vowel adjacency, it is typically the word-final vowel that becomes suppressed. Second, there is no substantial evidence for the lengthened free variants in the 1sg/3sg. If anything, the lengthened forms are the result of adjacency to word-initial nasal consonants whereby 1sg adopts its point of articulation.

Valkhoff made an important contribution by introducing an independent class of pronouns that occur in what he labels the disjunctive position and after prepositions.

Table 3. Disjunctive pronouns in Valkhoff (1966: 96).

	<i>1sg</i>	<i>2sg</i>	<i>3sg</i>	<i>1pl</i>	<i>2pl</i>	<i>3pl</i>
Disj./after prep. allomorphs free variants	<i>ami</i> <i>mi</i>	<i>bo</i>	<i>ele</i> <i>le ~ e</i> <i>elé</i>	<i>nõ</i>	<i>inãsé</i> <i>nãsé</i>	<i>ine</i> <i>ne</i> <i>iné</i>

This class of pronouns identifies strong forms that occur in specific environments that will be identified throughout this chapter. Valkhoff distinguishes between “emphatic” and “non-emphatic”, and in the case of 1sg *n* ‘I’ employs the term “prefixed” pronoun. The use of this term is perhaps related to his claim that 1sg, as well as 2pl and 3pl, are “of Bantu origin”. Valkhoff also claims that strong 3sg, i.e. *êlê*, occurs after prepositions. This finding can hardly be correct because in Negreiros (1895), Ferraz (1979) and modern Santome, weak 3sg *ê* is consistently used in this position. Furthermore, the same reasoning (as applied to *mi* in Table 2) extends to *mi* and *ele* in Table 3. Finally, 3pl in Table 3 should include a counterpart with a final nasalization (*inê*) and this also goes for the subject 3pl in Table 2. This is arguably the basic form, whereas all the other forms should be considered phonological variants.

2.2.3. Ferraz (1979)

Ferraz generally stands out for the quality of his data and provides a reliable description of Santome’s pronominal system. He introduces a distinction similar to the one found in Valkhoff (1966). His terminology differs however in that he refers to “bound” (e.g. 1sg *n ~ î*, and “free” morphemes (e.g., 1sg *ami*, 3sg *êlê*, 1pl *non*, etc.).

Table 4. Subject pronouns in Ferraz (1979: 62).

	<i>Singular</i>	<i>Plural</i>
1	<i>n ~ î, Ø, ami</i>	<i>nõ</i>
2	<i>bo</i>	<i>inãse ~ nãse</i>
3	<i>e, ele, Ø</i>	<i>inê ~ nê ~ ine</i>
Unmarked	<i>a</i>	

Ferraz identifies *ami* and *ele* as free forms that occur before “non-verbal forms” (or “other parts of speech”), corresponding obviously to strong forms. As for weak 1sg, Ferraz points out correctly that bound *N-* can be represented phonetically as [m, n, ŋ]. However, I was unable to find instances of 1st person *î* – the free variant of *n* in contemporary Santome.⁶ Furthermore, Ferraz is the first author to refer explicitly to what he calls the unmarked pronoun *a*, but did not identify all the environments in which this pronoun occurs (cf. section 2.3.2.).

The null-variant in his table corresponds to phonological cliticization and consequent incorporation of the nasal into a following nasal sound. In these cases, the 1st person pronoun adopts the articulation point of this nasal (e.g., **N mêsê* > *M mêsê* > *Mêsê* ‘I want’), although it should be noted that the nasal may become slightly lengthened in these conditions (cf. Table 2).

The null-variant of 3sg, Ferraz argues, is only licensed under previous identification, for instance in enumerations (*Ê fla*, [-] *kanta...* ‘He spoke, sang...’). This type of omission is of course common cross-linguistically. It is not fully clear to me why Ferraz included the null form for 3sg only. Provided the right syntactic conditions are met, other pronouns can also be elided under identification (cf. section 2.9.2.6). For this reason, the inclusion of null forms should arguably not be part of the paradigm.

Ferraz further shows that *ê* fills in the non-referential subject position, which will be discussed in section 2.9. Finally, he argues that the forms for 2pl and 3pl are free variants, a claim that in my view can only be applied to 3pl.

2.2.4. Schang (2000)

In his work on the emergence of the GGC, Schang (2000: 224-244) compares several aspects of personal pronouns in Santome, Ngola and Lung’ie. His results are similar to those found in Ferraz (1979), and are resumed in the following table:

⁶ Philippe Maurer [p.c.] informs me that this form is (still) used in Lung’ie, one of Santome’s sister creoles.

Table 5. Subject pronouns in Schang (2000: 225)

	<i>Singular</i>	<i>Plural</i>
1	<i>n-</i> <i>ami (disjoint)</i>	<i>nõ</i>
2	<i>bo</i>	<i>(i)nãse</i>
3	<i>e</i>	<i>ine /inẽ</i>
unmarked	<i>a</i>	

Like Valkhoff (1966), Schang adopts the term ‘disjoint’ for pronouns that occupy a non-argumental position (e.g. topics). The absence of strong 3sg *êlê* in his work is surprising, since this form is still fully productive. Schang claims that, contrary to Ferraz (1979), he found few occurrences of *êlê*. His informants considered this a form borrowed from Portuguese. My corpus reveals, however, that *êlê* is still extensively used as a strong subject, object and oblique pronoun and is obligatory in almost every environment it occurs in (cf. section 2.4). It should also be noted that *êlê* is not a recent form, since it can be found as early as the 20ties, in journal columns written by Francisco Bonfim.

Schang further proposes a split in four classes for subject, object and disjoint pronouns in which nasal 1sg *n* is treated as a clitic form (cf. Valkhoff 1966). He mentions that two of the pronouns are not truly personal, namely 3sg expletive *ê* (used in impersonal constructions) and *a* (cf. Ferraz). I will show in section 2.3.2 that the latter form actually does have pronominal reference in specific environments. Finally, Schang assumes with Ferraz that there are no disjoint plural forms. While this is indeed correct with respect to 1pl and 2pl, this assumption – as mentioned above - proves incorrect with respect to 3pl.

2.3. A new account of subject pronouns

Table 6 below offers a descriptive inventory of the subject pronouns I found in my naturalistic data. This table does not make any assumptions about the status of the pronouns. Section 2.4 to 2.7 focuses on the properties of the pronominal forms. The findings of these sections underlie the analysis proposed in section 2.8.

Table 6. Subject pronouns in Santome.

	<i>Singular</i>	<i>Plural</i>
1	<i>n, am, ami</i>	<i>non</i>
2	<i>ô, bô</i>	<i>a, nansê, nansê</i>
3	<i>ê, êlê</i>	<i>a, nen, inen</i>
2/3 respectful	<i>sun, sumu, sungê, sumungê</i> (masc.) <i>san, samu, sangê, samungê</i> (fem.)	<i>(i)nen sun/sumu/sungê/sumungê</i> (masc.) <i>(i)nen</i> <i>san/samu/sangê/samungê</i> (fem.)
Impersonal	<i>a</i>	
Expletive	<i>ê, kwa</i>	

In the next sections, I will briefly focus on several forms that have either been ignored in the literature on the pronominal system or have not been exhaustively described. For a good description of the pronominal paradigms, I refer the reader to Ferraz (1979:63-71).

2.3.1. Respectful pronouns

The forms of respect for 2sg/pl and 3sg/pl are unique compared to the other pronouns in the sense that they bear a gender distinction. Moreover, these forms also have the non-pronominal meaning ‘man, mister’ (*sun, sumu, sungê, sumungê*)⁷ and ‘woman, lady’ (*san, samu, sangê, samungê*). *Sun* and *san* are typically used anaphorically with either the same form or any of the other forms. This behavior is identical to that of a regular pronoun, as follows from comparing the examples (1) and (2), with the respectful forms, to (3) and (4), with standard pronouns.

- (1) *Sun na mêsê pa sun be ku mosu se fô?*
2SG NEG want for 2SG go with boy SP NEG-EMPH
‘Don’t you want to go with the boy?’
- (2) *Sel’ô sama san pa san bi kume.*
must-2SG call 3SG for 3SG come eat
‘You must call her to come to eat.’
- (3) *Non na mêsê pa non be ku mosu se fô?*
‘Don’t we want to go with the boy?’

⁷ The lexicalized forms *sumu* and *samu* arguably derive respectively from *sun mu* (*sun+my*) and *san mu* (*san+my*).

- (4) *Sel'ô sam'e pa ê bi kume.*
 must-2SG call-3SG for 3SG come eat
 'You must call her to come to eat.'

However, it is shown in examples (5-6) that *sumu* can be used as the antecedent and anaphor as well, provided there is the usual number and gender agreement. The same properties apply to the female counterpart *samu*.

- (5) *Sumu mêsê pa sumu/sun be.*
 2/3SG want for 2/3SG go
 'You want to go.' / He wants to go.'
- (6) *Sun ka lembla ku sumu so butxiza mu?*
 2SG ASP remember that 2SG FOC baptize me
 'Do you remember that you baptized me?'

Note that *sumu* also co-occurs with *sun* when used to express the highest degree of respect, usually for high-ranking persons in the society (kings in folk tales, priests, etc.), but this form can only be used as nominal modifier and not as a pronominal:

- (7) *Sun bi punta sumu sun padê xi manu ku mana ka kaza.*
 3SG come ask excellency priest if brother with sister ASP marry
 'He came to ask his excellency the priest whether the boy will marry her sister.'

Finally, the following examples show the use of *sungê/sangê* and *sumungê/samungê*.⁸

- (8) a. *Sungê ten ba ke sun.*
 3SG also go house POS
 'He also went to his house.'

⁸ *Sungê* and *sangê* are the contracted forms of *sun/san* and *ngê* 'person', but I assume these forms have become lexicalized. The same goes for *sumungê* and *samungê*, which is the contracted form of *sun/san+mu+ngê* (*sun/san+my+person*). The lexicalization of these forms follows for instance from the placement of specific marker *se*, a clitic which cannot intervene between the two forms:

- (i) a. *sun ngê se*
 man person SP
 b. **sun se ngê*
 'the man in question'

- b. *Ola ku ê glita, samungê punt'e...*
 when KU⁹ 3SG scream 3SG ask-3SG
 'When he screamed, she asked him...'

Unlike *sun/san* and *sumu/samu*, the forms *sungê/sangê* and *sumungê/samungê* cannot be used anaphorically.

2.3.2. Unmarked/anaphoric *a*

Here I will briefly address the pronominal form *a* (Ferraz 1979:66). This pronoun occurs in all the GGCs and has deserved special attention in the case of Ngola (Maurer 1995, Lorenzino 1996). Descriptively speaking, *a* in Santome is used as an impersonal pronoun, 'unmarked' in the terminology of Ferraz (1979), and as an anaphoric pronoun, a label first used by Lorenzino (1996). Examples (9-10) show the impersonal use of *a*.¹⁰

- (9) *Punda kamanda ku a sama nala Ototo?*
 Why why KU IMP call there Ototo
 'Why is it called Ototo over there?'
- (10) *A bamu fla santome.*¹¹
 IMP go speak santome
 'Let's speak Santome.'

In its anaphoric use, pronoun *a* is able to report back to an already specified 3pl referent. It can also occur in a 2pl subject-doubling construction.¹²

- (11) *So nen bixi kwa muntu ben. A ba misa.*
 then 3PL dress thing very well / ANAPH go mess
 'Then they dressed up nicely and went to mass.'

⁹ Since more research is necessary on the status of *ku*, I don not provide a translation nor a label for this item. In the literature on other creole languages, similar items have been analyzed as relative markers, focus markers and agreement markers.

¹⁰ Note that 2sg can also be used impersonally:

- (i) *Bô pô txila djêlu tlega mwala.*
 2SG may take money give woman
 'One may give money to his wife.'

¹¹ This sentence corresponds to the title of Alcântara (Ms.).

¹² According to Maurer (1995) and Lorenzino (1996), in Ngola *a* is also used for 2sg respectful.

- (12) *Sun Alê ku mwala sun saka vivê. A na tê mina fa.*
 Mr. King with woman POS ASP live / ANAPH NEG have child NEG
 ‘The king was living with his wife. They didn’t have children.’
- (13) *Inansê, a tava ba fesa.*
 2PL ANAPH TNS go party
 ‘You, you had gone to the party.’

Note that often *a* is also used in imperatives with plural reference (see also example (10) above).

- (14) *Inen mosu, a xa saku!*
 3pl boy ANAPH fill bag
 ‘You boys, fill the bags.’
- (15) *Pôvô ê, a ka fika ku Dêsu!*
 People EMPH ANAPH ASP remain with God
 ‘People, stay with God.’

Syntactically, *a* shares properties with weak 1sg *n*, since it cannot occur in environments where strong pronouns typically occur, except before *ten/tembeten* ‘also’, which is one of the environments that is also able to host weak 1sg/3sg (see section 2.4.3).

- (16) *A ten/*tudaxi fla kuma...*
 IMP also/also say that
 ‘It has also been said that...’

Given its properties, *a* only functions as a weak pronoun. Unlike other pronouns in the paradigm, it occurs exclusively in subject position and has to be anaphorically licensed when not used impersonally.

2.3.3. Special uses of third singular *ê*

In addition to the standard 3rd person reference, pronoun *ê* also occurs in several specific environments, which shows that it is used as a default or unmarked form. This is evident in the following sentences:

- (17) [*Êlê ku inen sode ku tudu ngê*] *be d'ê* *palaxu*.
 3SG with 3SG soldier with all people go PSR palace
 'He and the soldiers and everybody went away to the palace.'
- (18) [*Inen poto se*], *mosu fis'*[*e*].
 3PL door SP boy close-3SG
 'The doors, the boy closed them.'¹³
- (19) [*Tudu kwa ka bi*], [*ê*] *sa d'ôlô*.
 All thing ASP come 3SG be of-gold
 'All the things that came, they were made of gold.'
- (20) [*Ami*] *so ome fla ku ê*.
 1SG FOC man speak with 3SG
 'The man spoke to ME.'
- (21) *San tê* [*inen mina nala san*] *ku ome ka pô fla ku ê*.
 she have 3PL children there POS REL man ASP can talk with 3SG
 'She has the children of hers there with whom the men can talk.'
- (22) *Ê sôbê*.¹⁴
 'It rained.'
- (23) *Ê fika mosu, mina, men mosu*.
 3SG remain boy girl mother boy
 'The boy, the girl and the boy's mother remained.'

Example (17) shows that a plural, human antecedent can be recovered by a 3sg pseudo-reflexive pronoun (PSR) (cf. section 2.6.3 for more details on pseudo-reflexivity). In examples (18-19), it is evident that the left-dislocated antecedent does not have same person and/or number features as 3sg *ê*, by which it is recovered. It therefore follows that both human and non-human antecedents can be recovered by 3sg *ê*. However, animacy does play a role. If the antecedent in non-quantified left-dislocations such as (18) were human, the pronoun would have to agree in number with the antecedent.

¹³ Notice that this translation is awkward in English, but it best captures the spirit of the creole sentence.

¹⁴ More typically, the subject position is filled by a cognate subject.

- (i) *Suba ska sôbê*.
 Rain ASP rain
 'It is raining.'

Examples (20-21) are instances of preposition stranding with an invariable spelled-out trace (*ê*).¹⁵ In (20), the focused pronoun shows that 3sg even occurs when a pronoun (*ami*) is extracted, which underlines the unmarked role of 3sg in these constructions. Example (21) is a case of relativization. As in other creole languages, 3pl in Santome is also used as a plural marker¹⁶ (cf. *inen sode, inen mina* above) and thus clearly shows that there is no number agreement between the antecedent and the spelled-out trace.

Finally, examples (22-23) show different cases of the expletive use of 3SG, with a meteorological verb and a raising verb respectively.

It follows that in all the constructions above weak 3sg can be defective with respect to person, number or referentiality.

2.3.4. Expletive *kwa*

Another item that has gone unnoticed in the literature on pronouns is the special use of *kwa*, which normally means ‘thing’.¹⁷ However, there a number of cases where it is shown that *kwa* has evolved into an expletive. Consider the following examples.

(24) *Kwa sa ska fe kalôlô nai.*

EXPL be ASP make heat here

‘It has been hot in here.’

(25) *Kwa sa Dêsu vede ô.*

EXPL be God truth EMPH

‘It’s the true God.’

(26) *Kwa sa plama ê.*

EXPL be morning EMPH

‘It is morning.’

¹⁵ For the particularities of these constructions in Santome and other Portuguese-based creole languages, I refer readers to Alexandre & Hagemeyer (2002).

¹⁶ For an outline of the DP in Santome, see Alexandre & Hagemeyer (in press).

¹⁷ Note that *kwa* has many meanings and uses. It can be employed as an interrogative pronoun, in enters into lexicalized chunks with deictic, quantificational and interrogative meanings, etc.

- | | | |
|-----|------------------------------|--------------------|
| (i) | a. <i>Kwa bô fe?</i> | ‘what did you do?’ |
| | b. <i>Kwa se</i> | ‘this, that’ |
| | c. <i>Tudu (inen) kwa se</i> | ‘everything’ |
| | d. <i>K(w)a manda?</i> | ‘why?’ |
| | e. <i>Kwa kume</i> | ‘food’ |

(27) *Kwa ka sôbê.*

EXPL ASP rain

‘It rains.’

It follows that this form is especially prone to occur in presentational environments, preceding copula *sa* (24-26). Note further that *kwa* occurs in a lexicalized construction *da ku kwa*, as in (28).

(28) *N ga da bô ku kwa.*

1SG ASP hit 2SG with *kwa*

‘I will hit you.’

In this construction, *kwa* functions as a special place-filler for the instrument, which is specified when present (e.g. ...*ku po* ‘with a stick’). *Da* typically heads a double-object construction, but in this case it behaves like a light verb that forms a complex predicate with *ku kwa*. In this sense, *kwa* is argumental and Case-marked by the preposition, but, like expletives, lacks lexical content.

2.4. Evidence for a strong/weak bias

The previous literature on Santome’s pronominal system was shown to be unanimous with respect to the distinction between weak and strong forms, based mainly on the properties of the 1sg and 3sg forms (cf. Table 6). In this section I will therefore focus on the distribution of 1sg *ami* and *n*, on the one hand, and 3sg *êlê* and *ê*, on the other. These forms are sensitive to a number of syntactic tests that have been proposed in the literature since Kayne (1975), and other language-specific tests to be discussed in the following sections.

2.4.1. DP coordination and disjunction

Strong pronouns obligatorily occur in conjoined or disjoined structures. Weak forms are banned in these constructions.

(29) {*ami/*n; êlê/*ê*} *ku Zon.*

‘I/He and Zon.’

(30) *Nê {ami/*n; êlê/*ê} nê Zon.*

‘Neither I/he nor Zon.’

(31) *{ami/*n; êlê/*ê} ô Zon.*

‘I/He or zon.’

Note further that pronouns have exactly the same distribution as full DPs in both subject position, as in (32), and object position, in (33).

(32) a. *Inen mwala se ku mosu se ba ple.*

3PL woman SP with boy SP go beach

‘The women and the boy went to the beach.’

b. *Inen ku ê ba ple.*

3PL with 3SG go beach

‘He and they went to the beach.’

(33) a. *Zon tava ka pya inen mwala se ku mosu se.*

Zon TNS ASP look 3PL woman SP with boy SP

‘Zon was looking at the women and the boy.’

b. *Zon tava ka pya inen ku ê.*

‘Zon was looking at them and him.’

Contrary to what occurs in Romance languages, for example, Santome object pronouns are syntactically static forms, i.e. they do not exhibit any type of syntactic movement, which is also a preliminary, yet inconclusive, diagnostic against analyzing these forms as syntactic clitics (see sections 2.5-2.8).

2.4.2. Isolation

Another classical test concerns isolation of the pronoun, for example in a question-answer structure such as (34).

(34) Q: *Kêngê fe kwa se?*

who do thing SP

‘Who did that?’

A: *{ami/*n; êlê/*ê}*

‘I/He (did).’

As illustrated, the strong form is also obligatory in this environment.

2.4.3. Focus

It is generally held that weak pronouns and clitics cannot be focused; this is indeed borne out by the Santome data. Weak forms cannot be phonologically focused:

- (35) *n/ê fe kwa se.*
1SG/3SG do thing SP
'I/He did it.' / '*I/HE did it.'

Syntactic focus constructions, on the other hand, typically trigger the strong form, as follows from (36).

- (36) {*ami/*n; êlê/*ê*} *so ka be.*
1SG/3SG FOC ASP go
'It is I/he who goes.'

The limited range of items that may intervene between the subject and the extended VP, i.e. the VP + TMA markers and NEG, are the long form of the focus marker, *soku*, relative pronoun *ku*, and the adverbs *tan* 'only', *me/plopi* '-self', *tudaxi, ten, tembeten*¹⁸ 'also, too, as well' and *nai* 'here' and *nala* 'there' used non-spatially. The use of some of these items with pronouns is illustrated in (37-39).

- (37) {*ami/*n; êlê/*ê*} *tan ka be.*
1SG/3SG only ASP go
'Only I go / Only he goes.'

- (38) {*ami/*n; êlê/*ê*} *me ka be.*
1SG/3SG -self ASP go
'I myself go / He himself goes.'

¹⁸ In Ferraz (1979: 64) also mentions *tembe* 'also', but this appears to be an archaic form.

- (39) *Ami nai sa axi.*
 1SG here be like-this
 ‘I here am like this.’

In the literature, items like *tan* or *me* have been treated as exclusive particles, *tudaxi/ten/tembeten* are additive particles, and *nai/nala* exhibit deictic properties (e.g. König 1991, Krifka 1999). I will discuss several of these particles in more detail in section 4.5.1.1.1. Other instances of syntactic focus confirm the use of strong pronouns. Example (40) below is an example of contrastive focus with negation marker *fa*.

- (40) *Bô ba ke, {ami/*n; êlê/*ê} fa.*
 2sg go house 1SG/3SG NEG
 ‘You went home, not me/he.’

Finally, it should be noted that there is one case where an intervening focal element does not obligatorily trigger strong 1sg/3sg, namely the adverb *ten* ‘also’ and its long variant *tembeten*.¹⁹

- (41) *Ê/Êlê ten lanka da son môlê.*
 3SG also fall give ground die
 ‘He also fell dead on the ground.’
- (42) *Klôpô dana, n tembeten dana.*
 body spoil 1SG also spoil
 ‘The body got spoiled and so did I.’

This exception is only operational with additive adverb *ten*, which has the long variant *tembeten*, and does not extend to *tudaxi* ‘also’, which has to be preceded by strong forms. However, *tudaxi* is not a normal inclusive adverb because it also functions as a nominal/modifier with the meaning ‘everything, all’ (cf. 3.2.1.2). I will return to the case of *ten/tembeten* in section 2.6.1.

¹⁹ It is important to note that the weak forms can precede other focus markers, but only when there is lexical material to their left in order to license phonological cliticization (section 2.4).

2.4.4. Apposition

Apposition has also been used as a test for uncovering strong pronouns, since weak pronouns and clitics do not support appositives (e.g. DeGraff, 1993; Veenstra, 1996). As expected, Santome responds positively to this test as well:

- (43) {*ami/*n; êlê/*ê*}, *bon ngê ku pasa, konsê familya nansê.*
1SG/3SG good person that surpass know family 2PL
'I/He, a very good person, know(s) their family.'

2.4.5. Wh-in-situ

Wh-elements in Santome are normally moved to the clause-initial position but there is a single interrogative structure with *bô* 'where is', in which the Wh-element is exclusively clause-final.²⁰

- (44) {*ami/*n; êlê/*ê*} *bô?*
1SG/3SG where
'Where am I/is he?'

This Wh-element is restricted to nominal constituents. Note that the categorical status of *bô* is uncertain. However, it is highly unlikely that we are dealing with a locative verb, since *bô* cannot be preceded by TMA-markers and requires strong pronouns, whereas verbs are typically preceded by the weak form (cf. *ê/*êlê kume* 'he ate').

2.4.6. Subject doubling

Subject doubling confirms the distribution of non-strong and strong pronouns, which typically follow the pattern strong-weak, as illustrated in examples (45) and (46).

- (45) ... *punda ami, n sa mo Raul Wagner.*
because 1 SG 1 SG be like R. W.
'... because me I am like Raul Wagner.'
- (46) *Êlê, ê ba fesa.*
3SG 3sg go party

²⁰ This test is proposed by Aboh (2004: 139) for Gungbe.

‘As for him, he went to the party.’

Note that subject doubling is not a pervasive feature of Santome²¹, thereby differentiating itself from Upper Guinea creoles. In section 2.7 I will argue that the strong pronouns in these environments are topics.

2.4.7. Exclusiveness constructions

These structures are introduced by *sela* ‘except for, only’, which reverts the polarity of the clause it modifies.²²

- (47) *Nê ûa ngê na bi fa, sela {ami/*n; êlê/*ê}*.
not one person NEG come NEG except 1sg/3sg
‘Nobody came, only I/he.’

2.4.8. Modification

Aboh (2004) shows that in Gungbe the capacity to modify by attaching a specific marker is evidence for the strong status of a pronoun. This test also carries over to Santome, where specific marker *se* is able to modify strong 1SG *ami*, as illustrated:

- (48) *Nê ami/*n se ku sa nai en*.
nor 1sg SP REL be here EMPH
‘Nor I who is here.’

Note that this feature applies to the whole pronominal paradigm, except for 3sg (**êlê se*). I have no explanation for this exception.

2.4.9. Summary

The data above abundantly underscore the split between strong pronouns and non-strong forms. This split has been correctly diagnosed in previous work, but had not been subjected to extensive syntactic testing.

²¹ Ferraz (1979: 64) also claims this is not a frequent property and that *ami* has "emphatic significance" in this construction. He does not make reference to doubling with other subject pronouns.

²² For an overview of additional functions of *sela*, see Ferraz (1979: 84-5).

2.5. Phonologically reduced pronouns

From Table 6 it followed that an account of the pronominal system of Santome is not complete without taking into account those cases where pronouns have become phonologically reduced. Ferraz (1979: 64) was the first to observe that “on rare occasions” strong 1sg *ami* becomes reduced to *am*.²³ However, this feature is not restricted to *am*, and far more common than Ferraz suggests. Overt reduction occurs in: 1sg *am*, 3pl *nen* and 2sg *ô*. Since the properties of the latter form are quite distinct from the former two, section 2.5.1 focuses exclusively on *ô* and section 2.5.2 on *am* and *nen*.

2.5.1. Second person singular *ô*

This section focuses on the distribution of reduced 2sg *ô*. Note in the first place that the reduction of *bô* to *ô* is not unexpected, since Santome exhibits other cases of non-voiced bilabial suppression, especially at word-level. The reduced 2sg form is a unique case in the pronominal paradigm because it cannot occur in sentence-initial position, as demonstrated in (49-c).²⁴

- (49) a. *bô/*o ba ke.*
‘You went home.’
- b. *bô/*o tava ka ba ke.*
2sg TNS ASP go home
‘You were going home.’
- c. *bô/*ô ku Zon.*
‘You and Zon.’

These examples show that in the absence of phonological material to the left, *ô* is not available, which shows that we are dealing with a phonological clitic. The following examples corroborate this claim.

- (50) a. *N mêsê p'ô ku Zon be.* ← (*pa bô*)
1sg want for-2sg with Zon go
‘I want you and Zon to go.’

²³ The example he provides is *am tẽ* ‘I also’.

²⁴ In closely-related Ngola, a similar rule applies. *Bô* typically occurs in clause-initial position, whereas the reduced form *ô* requires a host to its left. Differently from Santome, however, the reduced form typically occurs in the remaining contexts, including strong positions (Maurer 1995).

- b. *Kamand'ô fe kwa se?* ← (*kamanda bô*)
 Why-2sg do thing SP
 'Why did you do that?'
- c. *Jina k'ô lembla...* ← (*ku bô*)
 since that-2SG remember
 'As far as you remember...'
- d. *Ol'ô ka xiga kinte...* ← (*ola bô*)
 when-2SG ASP arrive garden
 'When you arrive at the garden...'
- e. *Sel'ô sama san pa san bi kume.* ← (*sela bô*)
 must-2SG call 3sg for 3sg come eat
 'You must call her to come to eat.'

The examples (50a-e) show that *ô* can cliticize to, respectively, complementizers, Wh-words, prepositions, nouns and high modals²⁵. In other words, in my corpus I found that the 2sg subject pronoun typically cliticizes to functional material.

Furthermore, phonological reduction of *bô* to *ô* also occurs in the object and possessive paradigm. In (51), the possessive pronoun cliticizes to the noun and in (52) the object pronoun cliticizes to the verb. Finally, examples (50a) above and (53) below show that reduced 2sg can also occur in a strong syntactic environment, namely DP coordination.

- (51) *Kw'ô.* ← (*kwa bô*)
 Thing-POS
 'Your thing.'
- (52) *N d'ô kwa se.* ← (*da bô*)
 1SG give-2SG thing SP
 'I gave you the thing in question.'
- (53) *N ga d'[ô ku ê] livlu se.*
 1SG ASP give-2SG and 3SG book SP
 'I give you and him the book in question.'

²⁵ In Chapter 2, section 3.8.2, I argue that the position of *sela* corresponds to a specific functional projection in the lower portion of the left-periphery.

Finally, it should be noticed that reduced 2sg in object position is only able to cliticize to verbs with a low vowel ending (a).

- (54) a. *kume bô* → **kum'ô* / **kume'ô* 'eat you'
 b. *mêsê bô* → **mês'ô* / **mêsê'ô* 'want you'
 c. *fili bô* → **fil'ô* / **fili'ô* 'wound you'
 d. *vôlô bô* → **vôl'ô* / **vôlô'ô* 'make you angry'
 f. *golo bô* → **gol'ô* / **golo'ô* 'search you'
 g. *dumu bô* → **dum'ô* / **dumu'ô* 'pound you'

In this section I have demonstrated that reduced 2sg *ô* is a phonological enclitic. This pronoun occurs in any both weak and strong syntactic contexts, provided the right phonological conditions are met.

2.5.2. First person singular *am* and third person plural *nen*

A different case of phonological reduction from the one described in the previous section concerns 1sg *am* and 3pl *nen*, which are the reduced forms of, respectively, *ami* and *inen*. In both cases, vowel deletion produces the reduction of a dissyllabic pronoun to a monosyllabic one. One of the main differences compared to the non-reduced strong pronouns is the fact the reduced forms cannot occur in isolation nor occur with host appositives, as illustrated in (55) and (56) respectively.

(55) Q: *Kêngê fe kwa se?*

who do thing SP

'Who did this?'

A: {*Ami/*Am; Inen/*Nen*}

'I.' / 'They.'

(56) {*Ami/*Am; Inen/*Nen*}, *migu Zon*, ...

1SG/3SG friend Zon

'I/They, friend(s) of Zon ...'

However, *am/nen* are able to occur in some of typically strong contexts discussed in section 2.4, for instance in DP coordination, in (57), before focal adverbs, in (58), before contrastive negation, in (59) or preceding Wh-in-situ, in (60). Although not

explicitly demonstrated here, note that the non-reduced forms may, of course, occur in these environments as well.

- (57) [*Am/nen ku Zon*]_φ.
 ‘I/They and Zon.’
- (58) [*Am/Nen so*]_φ *ka be*.
 1SG/3PL FOC ASP go
 ‘It is me/them who goes/go.’
- (59) *Ê ba ke*, [*am/nen fa*]_φ.
 3SG go house 1SG/3PL NEG
 ‘He went home, not me/them.’
- (60) [*Am/Nen bô*]_φ?
 1SG/3PL where
 ‘Where am I/are they?’

Unlike in the case of apposition and isolation, in (55-56) *am* and *nen* are able to form a prosodic domain with the material to their right. However, the fact that in the case of *inen>nen* the left-edge of the pronoun is reduced, whereas the case of *ami>am* involves the right-edge, has implications for phonological phrasing. This becomes evident in the following examples:

- (61) *Zon [ku nen/ami/*am]*_φ
 ‘Zon and they/I.’
- (62) [_{DP} *Zon [ku nen/ami/*am]*]_φ *ba poson*
 Zon with 3PL/1SG go town
 ‘Zon and they/I went to the city of S. Tomé.’

From these examples it can be concluded that *nen* also has the ability to cliticize to the left, whereas this direction is precluded for *am* (*ami* must occur in these examples) Despite the fact that there is material to the right of the pronoun in (62), a syntactic frontier coinciding with a prosodic domain occurs to the left of the pronoun. Therefore *am* cannot cliticize to the verb *ba* and is banned from this position. In sum, *nen* is underspecified for the directionality of cliticization whereas *am* forms looks for a prosodic host to its right.

Another environment where *am* and *nen* show divergent behavior is subject doubling. Consider the following examples.

(63) *Ami, n ba fesa.*
1SG 1SG go party
'As to me, I went to the party.'

(64) *Inen, nen ba fesa.*
3pl 3pl go party
'As to them, they went to the party.'

It is shown in (64) that *nen* can function as the weak form in cases of subject doubling, whereas in (63) the strong form *ami* has to be doubled by weak *n*. This is expected in light of the more refined 1sg paradigm. In fact, abstracting of course away from phonological restrictions, the fact that both *inen* and *nen* occur quite freely in both weak and strong syntactic contexts differs from *ami* and *am*, which are generally considered slightly awkward in weak syntactic contexts, i.e. adjacent to the verb or TMA-material (cf. section 2.6.1).

Despite the difference between *am* and *nen* argued for in the preceding paragraph, the ability to be modified confirms the strong status of these two pronouns. *Nen*, being a plural form, can, for instance, be modified by numerals and quantifiers, in (65-a-b), whereas *am* can, for example, receive specific marking, in (66) (cf. also section 2.8.1).

(65) a. *Nen dôsu fla.*
3SG two speak
'The two of them spoke.'

b. *Tudu nen fla.*
All 3PL speak
'All of them spoke.'

(66) *Am se.*
1SG SP
'(The) I in question.'

The properties of strong reduced pronouns that were addressed in this section and in section 2.5.1 lead to the following conclusions:

- strong reduced forms have internal structure and thus clearly differ from the paradigm of weak pronouns.
- *inen* and *nen* cannot be considered free variants, as claimed by Ferraz (1979) and Schang (2000);
- the direction of phonological cliticization is not uniform: 2sg *ô* cliticizes exclusively to the left, 1sg *am* cliticizes exclusively to the right, and *nen* was shown to be underspecified;
- *nen* can occur in weak and strong contexts, whereas the use of *am* is restricted to strong environments.

I end this section with a brief examination of the 2pl forms *inansê/nansê*. Despite the existence of two forms that are reminiscent of the pair *inen/nen*, it should be noted that these forms of the 2pl pronouns are free variants, as claimed in earlier work. None of the restrictions detected in this section with respect to the occurrence of reduced strong 1sg and 3pl apply to these pronouns. There is evidence suggesting that the shorter *nansê* is in the process of replacing the longer form *inansê*. In my corpus, the occurrences of *nansê* clearly outnumber those of *inansê*.

2.6. Weak pronouns: syntax or phonology?

In section 2.4 it was shown that Santome exhibits a clear-cut distinction between the syntactic positions for weak and strong pronouns. However, in section 2.4.1 it was briefly mentioned that weak pronouns always occur in exactly the same environments as DPs. This also applies to object pronouns (section 2.6.2). In these circumstances, it is not immediately clear whether weak pronouns in Santome are a special type of syntactic clitic, namely agreement markers, or phonological clitics, like the pronouns discussed in section 2.5. The two hypotheses will be the object of study in the following sections.

2.6.1. Subject pronouns

In this section I will focus on 1sg and 3sg, because of the clear contrast between weak and strong forms (cf. section 2.4). It will be demonstrated that the properties of the

weak forms call for a more fine-grained distinction than what was suggested in section 2.4.

First, consider the equative constructions and the negative clauses, in respectively (67) and (68).

- (67) a. *n/ami sa dôtôlô.*
 1SG be doctor
 ‘I am a doctor.’
- b. *ê/*êlê sa dôtôlô.*
 ‘He is a doctor.’
- (68) a. *n/ami na ka kume pixi fa.*
 1SG NEG ASP eat fish NEG
 ‘I do not eat fish.’
- b. *ê/*êlê na ka kume pixi fa.*
 ‘He doesn’t eat fish.’

In these environments, both weak and strong 1sg are allowed. Strong 3sg *êlê*, however, is ungrammatical, as follows from the contrast between the a. and the b. sentences. A similar contrast between 1sg and 3sg applies when these pronouns immediately precede the verb, in (69), or core TMA-markers, in (70).

- (69) a. *n/?ami kume pixi.*
 ‘I/he ate fish.’
- b. *ê/*êlê kume pixi.*
 ‘He ate fish.’
- (70) a. *n/?ami {tava/ka/ska} kume pixi.*
 1SG/3SG TNS/ASP/ASP eat fish
 ‘I had eaten fish.’ / ‘I eat fish.’ / ‘I’m eating fish.’
- b. *ê/*êlê {tava/ka/ska} kume pixi.*
 ‘He had eaten fish.’ / ‘He eats fish.’ / ‘He’s eating fish.’

Note that unlike in (67a) and (68a) above, the use of strong 1sg *ami* in these environments is considered slightly awkward, showing speaker-dependent variation to some extent. My corpus of spontaneous speech in fact exhibits a very limited number of instances of strong 1sg in these syntactic positions. In fact, Ferraz (1979: 64) had

already noticed that *ami* in these positions corresponds to an “archaic usage”. Nevertheless, there is a clear-cut contrast between *ami* and *êlê*, since the latter shows a strong tendency to avoid standard Case-marking positions, favoring peripheral environments.

Equative clauses are presentational clauses, which often show differences in root and embedded environments. However, note that embedding the equative clause in (67) under finite complementizer *kuma* and non-finite complementizer *pa* does not trigger a difference in the selection of the pronominal form: both weak and strong 1sg can occur, in (71a) and (72a), whereas only weak 3sg is allowed in (71b) and (72b).²⁶

(71) a. *Bô sêbê kuma n/ami sa dôtôlô.*

2SG know that 1SG be doctor

‘You know that I’m a doctor.’

b. *Bô sêbê kum’ ê/*êlê sa dôtôlô.*

‘You know that he’s a doctor.’

(72) a. *Bô mêsê pa n/ami sa dôtôlô.*

2SG know for 1SG be doctor

‘You want me to be a doctor.’

b. *Bô mêsê p’ê/*êlê sa dôtôlô.*

‘You want him to be a doctor.’

Embedding typically strong environments, such as DP coordination and a Focus construction with adverb *tan*, reveals an important difference between weak 1sg and weak 3sg:

(73) a. *Bô fla pa [ami/am/*n ku Zon] ba ke.*

2SG tell for 1SG with Zon go home

‘You told me/ and John to go home.’

b. *Bô fla p’ [ê/êlê ku Zon] ba ke.*²⁷

‘You told him and Zon to go home.’

(74) a. *Bô fla p’[ami/am/*n tan] ba ke.*

²⁶ Here I use the terms finite and non-finite in the sense that in the former construction tense marker *tava* can occur in the embedded clause, while it is precluded in the latter environment.

²⁷ Note that in these cases the pronouns typically contract with the complementizer: *pa ê(lê) > p’ê(lê)* ‘for him/her’.

2SG tell for-1SG only go home

‘You told only me to go home.’

b. *Bô fla pa [ê/êlê tan] ba ke.*

‘You told only him to go home.’

These examples show that weak 3sg *ê* does occur in strong embedded syntactic environments, whereas 1sg *n* cannot. Since it was demonstrated in section 2.4 that weak 3sg could not occur in an ‘out-of-the-blue’ DP coordination (*Êlê/*ê ku Zon* ‘He and Zon’), it follows that the phonological environment in (73b) and (74b) above rescues this pronoun. These findings are confirmed by the fact that weak 3sg can occur as the object of a preposition, a typical XP position, whereas weak 1sg cannot.

(75) *Zon ku ê/*n.*

‘Zon and he/I.’

It follows that *ê* cliticizes phonologically to the left. As Déprez (1994) points out, this is rather unexpected under an account that analyzes subject clitics as agreement markers in T (INFL in her proposal), because it would imply, against the facts, that there is head movement from T to C.

Déprez also argues that adverb placement between pronouns and following material is problematic for the syntactic analysis. As illustrated in (76), weak pronominal forms in Santome can precede adverb *ten* ‘also’ (cf. also Ferraz 1979: 63):

(76) *N/Ê ten tê ngê nala.*

1SG/3SG also have people there

‘I/He also have/has people (I know/he knows) there.’

It should be noticed that this is a productive construction and native speakers consider it fully grammatical. In section 2.4.3 it was shown that other focal adverbs that surface arguably in the same position as *ten* require the strong form. In the case of 3sg *ê*, the occurrence of the pronoun in this position adds up to several other facts that are incompatible with a syntactic cliticization. *Ê* is thus a weak syntactic pronoun, not a clitic. As for 1sg *n*, things are quite different, since the adverb placement test above is the first and, as it turns out, the only potential argument against a syntactic analysis.

In 4.5.1.1.1, I argue that focus adverbs such as *tan* ‘only’ are adjoined to DP and take scope to the left. These adverbs always require a strong pronoun. *Ten/tembeten*, however, is not only different because of the possibility that it may be preceded by weak pronouns, it also exhibits two scope directions: narrow scope over the subject or wide scope over the predicate. Therefore, I have explored the hypothesis that weak and strong pronouns preceding *ten/tembeten* can be explained according to the scope of the adverb. However, according to my informants, both directions of scope are equally possible with the weak and the strong pronoun.

- (77) *n/ami ten kume pixi.*
 1SG also eat fish
 ‘I too ate fish.’
 ‘I ate fish too.’

Hence, scopal properties of the adverb are not responsible for triggering the weak and the strong form. In light of (78), where *n* precedes and follows the adverb, it must be the case that *n+ten* occur high in the structure. It is suggestive that the highest *n* is in the same position as *ami* in (79).

- (78) *N ten n ga sa ke.*
 1sg also 1SG ASP be house
 ‘As to me, I’ll stay at home.’
- (79) *Mbon, ami ten n na sêbê nadaxi fa ê.*
 well 1SG also 1SG NEG know nothing NEG EMPH
 Well, as to me, I don’t know anything.’

Hence, independently of the details of a syntactic analysis for these cases, a case cannot be made that *n* is an agreement marker in (78), for instance by analyzing *ten* as an X° that adjoins to T° . Therefore, I will use this construction as a unique yet decisive piece of evidence against the hypothesis that weak pronouns are syntactic clitics.

Finally, two other minor arguments in support of a phonological approach are available. Since Schuchardt (1882), it has been claimed by most authors that *n* is a bound form. In fact, it can readily be shown that this pronoun cliticizes to the right.

Ferraz shows that *n* adopts the point of articulation of the initial consonant of a verb. This can be seen in the following contexts:

- (80) a. [**n**] *tlaba* ‘I worked’ (Ferraz 1979: 63)
b. [**m**] *bi* ‘I came’
c. [**N**] *gosta* ‘I liked’

Note also that Santome exhibits a phonologically conditioned variant of aspect marker *ka*, which becomes *ga* when preceded by 1sg *n*.

- (81) *So* [**ŋ**] *ga fla...*
then 1SG ASP speak
‘Then I say.’

Assuming that *n* in these cases cliticizes to the right, as usual, the aspect marker’s initial consonant becomes voiced through the nasal. Voicing of the aspect marker does not take place when it is preceded by, for instance, 1pl *non* or 3pl *inen*. I therefore consider this feature an idiosyncrasy of the language, comparable with, for instance, the pronunciation of *je suis* as [ʃu] in French copulative environments. Labelle (1985) argues that these idiosyncrasies constitute evidence for a phonological analysis.

A final argument in support of a generalized phonological analysis relates to the fact that the weak pronouns are presumably derived from the strong pronouns:

- (82) a. 1sg *ami* > *am* > **m* > *n*
b. 2sg *bô* > *ô*
c. 3sg *êlê* > **êl* > *ê*
d. 3pl *inen* > *nen*

Languages with true syntactic clitics generally have at least some unrelated strong and clitic forms for the same person (e.g. Lefebvre 1998).

2.6.2. Object pronouns

One of the arguments used by Déprez (1994) in her argumentation against syntactic clitics in Haitian is the behavior of object clitics in that language. Object positions corroborate that pronouns may cliticize to the left but without being related to verb movement. In this respect, a parallel can be established between Santome and Haitian. Note in the first place that the subject and object paradigm are virtually identical in form.

Table 7. Direct object pronouns in Santome.²⁸

	<i>Singular</i>	<i>Plural</i>
1	<i>m, mu, mun</i>	<i>non</i>
2	<i>ô, bô</i>	<i>inansê ~ nansê</i>
3	<i>ê²⁹, (êlê)</i>	<i>inen ~ nen ~ ne</i>
2/3 respectful	<i>sun, sumu (masc.)</i> <i>san, samu (fem.)</i>	<i>(i)nen sun/sumu (masc.)</i> <i>(i)nen san/ samu (fem.)</i>

By comparing Table 6 and 7, it follows that only the 1sg object pronoun has a different realization compared to the 1sg subject pronoun. To the best of my knowledge, object pronoun *êlê* only occurs with the verb *bê* ‘to see’, which is why I present it in parentheses. As will be shown below, this is an idiosyncratic construction.

The strong adjacency requirement of object pronouns with respect to the verb cannot be considered a valid argument in support of syntactic cliticization, since the verb remains *in situ* in Santome (3.2.1). As in most creole languages, this implies that no lexical material can intervene between the verb and the object (cf. Roberts 1999).

Apart from the abovementioned case of strong 3sg with *bê*, strong forms cannot occur in object position, as shown by the following examples:

- (83) a. *Bô sama mu/*ami.*
 2sg call 1SG
 ‘You called me.’

²⁸ Cf. also Ferraz (1979: 62)

²⁹ Realized as [e] or [ɛ] in the relevant contexts.

- b. *Bô sam'e/*êlê.*
 2sg call-3SG
 'You called him.'

Whenever cliticization on the verb is impossible, as in the specific case of double object constructions, the strong form is required, as illustrated in (84). Note further that the goal argument always precedes the theme in these cases (section 3.2.1.3).

- (84) *Bô da mu êlê/*ê.*
 2sg give 1SG 3SG
 'You gave it to me.'

When the verb selects a coordinated complement, the judgments of my informants are not always uniform. It stands out that in general both the strong and the weak form can occur in this type of environment, as illustrated:

- (85) a. *Zon {zuda-ami/?mu; zud-êlê/?e} ku Maya.*
 Zon {help-1SG; help-3SG} with Zon
 'Zon helped me/him and Maya.'
- b. *Zon {golo-mu/?ami; golo-e/?êlê} ku Maya.*
 Zon {search-1SG; search-3SG} and Maya
 'Zon searched for me and Maya.'
- c. *Zon {fili-mu/??ami; fili-ê/??êlê} ku Maya.*
 Zon {wound-1SG; wound-3SG} and Maya
 'Zon wounded me and Maya.'

The variable judgments are arguably related to phonological rules. In any case, weak forms can occur in this position and are sometimes even obligatory (*fili*), thereby differing from embedded clauses such as in (73a) and (74a), where it was shown that weak 1sg could not occur when a syntactic context for strong pronouns was embedded. Assuming that 1sg subject pronoun *n* and 1sg object pronoun *mu* are equally weak, this contrast can arguably be explained in terms of locality: verbs and objects are strictly adjacent sisters under the same maximal projection (VP); complementizers and subjects belong to two different maximal projections, arguably CP and TP, and are not strictly

adjacent, since lexical material such as adverbs may intervene. In fact, phonological cliticization is sensitive to phonological phrasing, which on its turn is sensitive to syntactic bracketing. It was shown that subject pronoun *n* cliticizes rightwards, whereas object pronouns cliticize leftwards. This is particularly clear in the case of weak 3sg, which undergoes different phonological processes, such as vowel harmony, semi-vocalization and contraction, as illustrated:

- (86) a. *da ê* → [dɛ] ‘give it’
 b. *kume ê* → [kumê] ‘eat it’
 c. *sêbê ê* → [sebe:] ‘know it’
 d. *ligi ê* → [ligi^ye] ‘lift it up’
 e. *dumu ê* → [dum^we] ‘pound it’
 f. *vôlô ê* → [volô^we] ‘be angry at him/her’
 g. *golo ê* → [gɔlɔ^wɛ] ‘look for it’

Hence, object pronouns confirm the importance of phonological cliticization in Santome, but also confirm that the behavior of subject 1sg *n* cannot be explained by phonology alone.

2.6.3. Pseudo-reflexive pronouns

This section focuses on the properties of pseudo-reflexives (PSR) as an additional piece of evidence for phonological cliticization in Santome.

Ferraz (1979: 72) claims that the only reflexivization strategy available in Santome is without overt reflexive marking, for instance:

- (87) *N ga ba kenta ∅.*
 1SG ASP go warm
 ‘I am going to warm myself.’

However, Ferraz fails to mention two strategies that equally encode reflexivization, namely so-called body-part reflexives with nominal *ubwê*, in (88) and pseudo-reflexivization, in (89).

(88) *N da ubwê mu ku faka.*

1SG give body POS with knife

‘I cut myself with a knife.’

(89) *N be mu fela.*

1SG go PSR market

‘I went away to the market.’

The paradigm of pseudo-reflexives is exactly the same as the possessive paradigm, as follows from Table 8.

Table 8. Possessive and pseudo-reflexive pronouns

	<i>Singular</i>	<i>Plural</i>
1	<i>mu</i>	<i>non</i>
2	<i>bô</i>	<i>nansê ~ dinansê</i>
3	<i>dê</i>	<i>nen~dinen</i>
2/3 respectful	<i>sun, sumu</i> (masc.) <i>san, samu</i> (fem.)	<i>(di)nen sun/sumu</i> (masc.) <i>(di)nen san/samu</i> (fem.)

Pseudo-reflexives can occur with a wide range of verbs of which I would like to highlight manner of motion verbs, in (90), directed motion verbs, in (91), and true unaccusatives, in (92).

(90) *Tudu xitu ku bwê ska kôlê dê...*

all place REL cow ASP run PSR

‘All the places where the cow is running around...’

(91) *N xê mu ni ke.*

1SG leave PSR from house

‘I left the house.’

(92) *Didi mu ô, n ga môlê mu ê.*

Didi my EMPH 1SG ASP die PSR EMPH

My dear Didi, I’m going to die.’

I label these forms pseudo-reflexives because they do not actually add a reflexive meaning to the clause. Spanish, for instance, exhibits these forms with some directed

motion verbs (e.g. Sp. *se fue* ‘has left/gone’) and some unaccusatives (Sp. *me muero* ‘I die’). Santome differs from Spanish in that pseudo-reflexives in the latter language are identical to the pronominal object paradigm and behave as syntactic clitics. I will argue that pseudo-reflexives in Santome are phonological clitics and underlying PPs. This analysis is based on the 3sg, 2pl and 3pl forms that have the following structure: *di* ‘of, from’+pronoun. This is evident with 3sg and 3pl forms, as in (93).

- (93) a. *Zon mole dê.*
 Zon die PSR (=of-3SG)
 ‘Zon has died.’
- b. *Inen be dinen.*
 3PL go PSR (= of-3PL)
 ‘They went away.’

I assume that pseudo-reflexives and PPs have essentially the same structure, and should be considered right-adjuncts to VP (section 3.2.1.4 and sub-sections). They differ in that only pseudo-reflexives exhibit the behavior of phonological clitics. First, consider the following examples:

- (94) a. *Zon be [PP dai losa].*
 Zon go from-here plantation
 ‘Zon went from here from the plantation.’
- b. *Zon ba [DP losa] [PP dai].*
 Zon go plantation from-here
 ‘Zon went to the plantation from here.’
- c. *Zon be [PP dai] [VP ba losa].*
 Zon go from-here go plantation
 ‘Zon went from here to the plantation.’
- (95) a. *Zon be [-]_i [PP dê] [DP losa]_i.*
 Zon go PSR plantation
 ‘Zon went away to the plantation.’
- b. *#Zon ba [DP losa] [PP dê].*

The goal-denoting argument *losa* ‘plantation’ is the primary argument of the verb *be/ba* ‘to go’. Note that the form *ba* can only be triggered when the goal DP is adjacent to the verb (cf. Hagemeijer 2000, 2005a). If we compare (94a) to (95a), it follows that the goal meaning cannot be obtained in the former example. The ‘from-to’ interpretation is, however, available in (94b), where the PP is arguably in its basic position (right-adjoined to VP) or in (94c) by adjoining a second VP and thereby serializing the construction. The pseudo-reflexive cannot occur after the goal.³⁰

Hence, I propose that in (95a) the PP *dê* is in its basic position, as is *dai* in (94b), but the goal argument has obligatorily extraposed to comply with the strict adjacency between the verb and the pseudo-reflexive. In an example like (94a), on the other hand, it follows from the interpretation that the locative (*losa*) is not extraposed, forming a complex PP with the PP.

In sum, pseudo-reflexives are externally merged as adjunct PPs with the status of weak pronouns for syntax and clitics for phonology. The goal of a directed motion verb, such as *ba/be* above, is the primary argument of the verb but has to be extraposed in order to allow cliticization of the weak pseudo-reflexive pronoun to the verb. In section 3.2.1.4, I will provide a more in-depth discussion of these constructions in my argumentation against verb movement. The demonstrable absence of verb movement is by itself an argument in support of a non-syntactic approach to pseudo-reflexives, in the sense that these pronouns are not generated in a CliticPhrase (e.g. Sportiche 1996).

2.6.4. Looking for a host

So far I have essentially argued for a phonological approach to weak pronouns in Santome. However, as mentioned on several occasions in this chapter, the authors that propose a syntactic account in creole languages generally lack evidence for truly syntactic properties of pronouns: I am not aware of any creole language that exhibits the type of clitic movement rule found in Romance languages. Nevertheless, this does not *a priori* invalidate the claim that creoles have or are developing a special type of syntactic agreement markers. This section will therefore investigate whether there is a unifying host that could potentially lend support to a syntactic analysis of weak pronouns in Santome.

³⁰ Note that, strictly speaking, (95b) is not ungrammatical if *dê* is interpreted as a possessive, ‘his/her plantation’.

Standard accounts of syntactic clitics typically posit two types: verbal clitics, as in modern Romance, and second-position clitics (Wackernagel position) following the first syntactic constituent or the first prosodic word, as in Serbo-Croatian.

However, several authors have abandoned the mainstream analyses by hypothesizing that syntactic clitics do not necessarily have specialized hosts. Fiéis (2003) shows that clitics in Old Portuguese were not only hosted by X° but also by XP (13th to 16th century). After moving through an exclusive X° stage (16th century), the host started becoming fully specified as V° , which corresponds to the situation in modern Portuguese. Several other scholars also entertain the possibility that clitics in Old Romance could adjoin to X° and XP (e.g. Rivero 1997, Duarte & Matos 2000). In Portuguese, the claim that at a certain stage XPs were able to host clitics is based on interpolation data, where lexical material like negation or adverbials intervene between the proclitic and the verb. Note that it has been suggested that in European Portuguese the negation marker and interpolated adverbs are verb-adjoined heads, because I-to-C raising shows that the whole complex moves (cf. Fiéis, 2003: 421-22).

As for the case of Santome, I have claimed that weak pronouns are best analyzed as phonological clitics. A purely syntactic account of pronominal cliticization runs into the basic problem that there is no unifying feature for the different types of hosts that pronouns can attach to, since they can immediately precede a large array of categories: the verb, any TMA-marker, negation and even inclusive adverb *ten~tembeten* ‘also’.

If subject pronouns were to be considered strictly verbal clitics, this would require that all the material that can intervene between the verb and subject pronouns has to be of clitic or affixal nature in order to form a clitic cluster with the pronoun. In recent work this hypothesis has been tested for the case of French (De Cat 2005). In addition to the more standard claim that French subject clitics are syntactic clitics (e.g. Kayne 1975, Rizzi 1986), it has sometimes been proposed in the literature that French subject clitics are morphological items that adjoin to the verb in the lexicon (e.g. Auger 1994). De Cat (2005) discusses both approaches, and argues against the morphological analysis on the grounds that French subject clitics are available for syntactic movement and that intervening lexical material between the verb and the clitic, more specifically the negation marker *ne*, genitive *en/y* and object clitics, cannot be treated as affixes.

If we look at the intervening elements in Santome, it follows that there are no compelling reasons to believe that they all behave like clitics or, even less plausible so,

affixes. This is particularly clear in the case of negation, tense marker *tava* and intervening adverbial material.

Standard negation marker *na* and its morphologically complex counterparts *naxi* ‘not yet’ and *nanta(n)* ‘not anymore’ (section 4.3) head a base-generated syntactic position on top of VP and the projections for TMA-material and there are no arguments to support that negation markers are clitics or affixes. Since there is no verb movement in Santome (cf. Chapter 3), this is very different from a language like French, where clitic *ne* is raised with the verb (Zanuttini 1997).

In section 3.3.6, I will argue that tense marker *tava* is base-generated as the head of TP in certain constructions, for instance in the following case:

(96) *Zon tava kwaji kume.*

Zon TNS almost eat

‘Zon had almost eaten.’/ ‘Zon was almost eating.’

Here I am interested in the pluperfect reading³¹, for which I assume that *tava* heads TP and that adverb *kwaji* is an adjunct merged to AspP, a functional projection that sits between TP and VP. It follows immediately that *tava* cannot be a clitic in this configuration, casting serious doubts on a syntactic analysis of weak pronouns.

A further complication is that one would have to assume that adverb *ten~tembeten* is a clitic. In fact, aspect markers are the only intervening lexical items that qualify as bound morphemes to the verb (section 3.3.5).

Nevertheless, the problem of the host for the syntactic analysis has not gone unnoticed, and led DeGraff to assume that preverbal TMA-markers isolate what corresponds to verbal inflection in morphologically richer languages. These TMA-markers often become grammaticalized verbs in creole languages. Thus, in this view, the unifying feature of the host would be a verbal feature. However, there is a measure of controversy to this if we apply this argument to aspect and mood markers in Santome. Even though TMA-markers belong to the extended VP projection in the sense of Grimshaw (1991), this assumption cannot be easily extended to the intervening adverb *ten*. Thus, the lack of a specialized host for subject pronouns further weakens the hypothesis that these forms are syntactic clitics in Santome. It can readily be shown that

³¹ For a detailed analysis of these interpretations, I refer the reader to Chapter, in particular section 3.7.

syntactic cliticization of weak pronouns imposes too heavy a burden on the analysis, whereas under a phonological approach the results are self-evident.

2.7. Strong pronouns as topics

This section will argue for strong pronouns as topics in Santome. I will examine the following three structural hypotheses for subject pronouns:

- (97) a. $[_{TP} XP [_T X^{\circ}]]$
b. $[_{TopP} XP [_{TP} XP [_T]]]$
c. $[_{TopP} XP [_{TP} pro [_T cl+X^{\circ}]]]$

Configuration (97a) is an instance of Spec-Head agreement and reflects the standard claims for subject clitics in, for instance, some Franco-Provençal and Italian dialects (e.g. Rizzi 1986, De Crousaz & Shlonsky, 2000).

In configuration (97b), there are no syntactic clitics, and strong pronouns are topics or, depending on information structure, something else, for instance focus. Standard subject pronouns are thus treated as XPs under this analysis. This is the configuration I will adopt for the facts in Santome.

Finally, (97c) is the type of structure that has, for instance, been proposed for Bantu languages (e.g. Bresnan & Mchombo 1987, Demuth & Gruber 1995). In Bantu, agreement affixes (X°) are generally obligatory and, as in Romance, the preverbal or postverbal position of the subject is bound to information structure.

Hypothesis (97c) has to be assessed given the putative role of Bantu in the formation of the GGC. Ferraz (1979), Maurer (1992) and Lorenzino (1998) show that western Bantu (Kikongo, Kimbundu) had a significant lexical impact on the GGC. Furthermore, in Palenquero (Colombia), where Kikongo appears to have had an almost exclusive substrate impact (Schwegler 2006), it seems likely that a very significant portion of Kikongo pronominal morphosyntax was transferred into the creole (Schwegler 2002).

The configuration in (97a) is highly problematic for several reasons. First, if this were the correct interpretation of the pronominal data, doubling would be expected to be a widespread feature of Santome, which it is not. Although contexts of subject doubling can easily be elicited, this structure is attested only sporadically in my corpus. Another serious drawback for the type of analysis in (97a) is this: by itself, strong 3sg *êlê* cannot

be the subject of root and embedded clauses; at the same time, *ami* was shown to only marginally occur as the subject of root clauses but not in embedded contexts.

Since (97a) is empirically inadequate, I will explore the idea that in the subject doubling structures strong 1sg/3sg should be analyzed as topics, as predicted under hypotheses (97b-c). The first argument that points in this direction is the strong intonational break between, for instance, strong and weak 3sg or 3pl. The fact that this intonational break is generally not (as) perceptible in the doubling sequence *ami n* can be readily explained if we take into account the phonological status of weak *n*.

The intuition that strong forms are topics finds support in the fact that 3sg *ê* cannot “double” a quantifier in subject position. To highlight this argument, consider the following examples of a Franco-Provençal dialect studied by De Crousaz & Shlonsky (2000):

- (98) a. *Djan (i) medzè na fondia mitya-mitya.* (Franco-Provençal of Gruyère)
 John (SCL) eats a fondue half-half
 ‘John is eating a ‘moitié-moitié’ fondue.’
- b. *Djan, *(i) vinyè.*
 ‘John, he came.’
- c. *Kôkon (i) vinyè.*
 ‘Somebody came.’

The grammaticality of the co-occurrence of a quantificational subject and a X° subject clitic, in (98a) shows that dislocation did not take place. The b. example is a hanging topic exhibiting an obligatory subject clitic. Quantificational subjects cannot be dislocated via a discourse pause, and exhibit an optional subject clitic as in (98c). Therefore *Djan* in (98a) and *Kôkon* in (98c) behave like regular subjects in a specifier position. As expected, in Santome quantificational subjects cannot be dislocated (by a break) nor be doubled by *ê*.

- (99) *Ûa ngê (*ê) bi.*
 a person (3SG) come
 ‘Somebody came.’

Thus, I assume that this subject is in the canonical subject position. It now becomes clear that strong 3sg is indeed a dislocated argument. Topics are concerned with “aboutness” and this is precisely the effect that subject doubling exhibits in Santome, as illustrated by the following sequences:

(100) *Kada ngê tê kwa dê. Ami n ga gôgô ku vungu.*
 Each person have thing POS. 1SG 1SG ASP like with music
 ‘Everybody has his thing. As for me, I like music.’

(101) *Tudu pletu di Santome ku Plinxi pi sa familya. Ami n sa familya F.*
 All black of S. Tomé and Príncipe be family. 1SG SG be family F.
 ‘All black people of São Tomé and Príncipe are family. As for me, I belong to the Fernandes family.’

Thus, whenever doubling takes place, the strong leftmost pronoun occurs structurally to the left of TP in a TopP.

Another drawback for the syntactic analysis is that hypothesis (97c) posits null subjects. This is unproblematic for the relevant Italian dialects and Bantu languages, which are characterized by rich verbal morphology to identify *pro* and by the possibility of spec-head agreement between the clitic and the non-clitic pronoun. Languages like Mandarin also stand the test because of its discourse-oriented properties. Since I have shown that the syntactic analysis could not be upheld for Santome and perhaps for other creole languages as well, in particular for Haitian (Déprez 1994), I will not adopt hypothesis (97c). Moreover, it should be noted that Santome exhibits hardly any non-argumental or pseudo-argumental *pro*-drop, as argued in section 2.9.

In the light of the empirical and cross-linguistic evidence currently under discussion, it follows that subject doubling in Santome best complies with the following representation.

(102) [_{TopP} XP [_{TP} XP [_{T'}]]]

The implication is therefore that Santome is not a typologically special kind of *pro*-drop language.

2.8. Pronouns at the syntax-phonology interface

This section provides a schematized overview of the properties of subject pronouns in Santome, on the one hand, and a discussion of previous theoretical claims about the workings of the pronominal system, on the other. Note that Table 9 does not include the respectful pronouns (cf. table 6 and section 2.3.1). The use of these forms is unrestricted, like 1pl *non* and 2pl *inansê*.

Table 9. Properties of subject pronouns.

	<i>Singular</i>							<i>Plural</i>					
	1			2		3		1	2		3		2/3
	<i>n</i>	<i>am</i>	<i>ami</i>	<i>ô</i>	<i>bô</i>	<i>ê</i>	<i>êlê</i>	<i>non</i>	<i>nansê</i>	<i>inansê</i>	<i>nen</i>	<i>inen</i>	<i>a</i>
Stress	*	✓	✓	*	✓	*	✓	✓	✓	✓	✓	✓	*
Focus I ³²	✓	✓	✓	*	✓	✓	✓	✓	✓	✓	✓	✓	✓
Focus II	*	✓	✓	*	✓	*	✓	✓	✓	✓	✓	✓	*
Coordination ³³	*	*	✓	*	✓	*	✓	✓	✓	✓	✓	✓	*
Isolation	*	*	✓	*	✓	*	✓	✓	✓	✓	*	✓	*
Apposition	*	*	✓	*	✓	*	✓	✓	✓	✓	*	✓	*
Wh-in-situ	*	✓	✓	*	✓	*	✓	✓	✓	✓	✓	✓	*
Modification ³⁴	*	✓	✓	*	✓	*	✓	✓	✓	✓	✓	✓	*
Negation	✓	✓	✓	*	✓	✓	*	✓	✓	✓	✓	✓	✓
TMA/Verb	✓	?	?	*	✓	✓	*	✓	✓	✓	✓	✓	✓

Based on the table, a number of generalizations can be made:

- (i) there is a weak/strong distinction that cuts through a number of pronouns;
- (ii) reduction and phonological cliticization is an important feature of this language;
- (iii) there are generally fine-grained differences between pronouns, even when they exhibit broadly the same properties.

Therefore, I propose the intersected classification of subject pronouns in Santome that is presented in Table 10:

³² Focus I stands for the possibility of the pronoun to precede adverb *ten* and *tembeten* ‘also’; Focus II stands for the possibility of the pronoun to precede all the remaining focus particles (cf. section 4.3).

³³ Coordination stands for the possibility of the pronoun to occur as the first term in DP-coordination.

³⁴ By ‘modification’ I mean that the pronoun may be modified by, for instance, numerals, specific marker or quantifiers. This test therefore concerns the internal structure of the pronoun.

Table 10. Distribution of subject and object pronouns at the syntax-phonology interface.

		<i>Syntax</i>		
		Weak	Strong	Underspecified
Phonology	+Clitic	<i>a, n, mu, ê_{expletive}</i>	<i>am</i>	<i>ô, nen, ê</i>
	-Clitic	<i>kwa</i>	<i>ami, êlê</i>	<i>bô, sun/san, non, (i)nansê, inen</i>

Since it was shown that subject and object pronouns in Santome are identical, except for 1sg, which is the most prominent pronoun in the pronominal hierarchy and syntactically exhibits a total of four different forms for the subject (*n, am, ami*) and object position (*mu*). It is stated that pronouns in Santome can either be weak, strong or underspecified with respect to syntax. I have adopted the label ‘underspecified’ in order to account for pronouns that occur in typically weak and strong positions³⁵, which are respectively the position before the verb and TMA-markers and peripheral position (Focus, Topic, etc.) in the case of subject pronouns. For object pronouns, the weak position is the position adjacent to the verb. Note that pronouns do not always present a clear-cut behavior with respect to syntax. *Ami* is a good example, because it was shown that it occurs typically in strong environments. Nevertheless most speakers still accept this pronoun preceding the verb and TMA-markers, although this is generally considered slightly marginal. Therefore, I have opted to include it in the strong class and not in the underspecified class. In the light of the available evidence, the strongest pronoun in the whole paradigm is 3sg *êlê*. Pronouns such as 2sg *bô*, 1pl *non* or 3pl *inen* occur unrestrictedly in weak and strong syntactic environments.

At the same time, the table implies that pronouns behave phonologically like clitic or non-clitic forms. Although I assume that pronouns can not be underspecified for this feature, it was shown in section 2.5.2 that they can be underspecified for the direction of phonological cliticization, which is a subparameter of [+/- phonological clitic]. There are of course many other subclassifications of pronouns that uncover fine-grained distinctions, such as the ability to bear a gender specification (*sun/san*) or to be referential. The latter specification distinguishes expletive *ê* and *kwa* from the remainder of the paradigm. Note further that these two expletives are weak syntactic forms,

³⁵ The inclusion in more than one class of pronouns is also suggested by, for instance, Cardinaletti & Starke (1995: 31) for German pronouns that are ambiguous between weak and strong or by Aboh (2004: 143) for 3sg in Gungbe, which he labels a ‘mutant form’, between a weak and a clitic pronoun.

because they cannot occur in the strong contexts discussed in section 2.4, except preceding *ten*. One of the features that has been discussed is the ability of a pronoun to refer to items that are [-human], a property that is arguably unavailable for strong pronouns (e.g. Cardinaletti & Starke 1995, Corver & Delfitto 1993). Santome challenges this assumption, since it was demonstrated that in the object paradigm strong 3sg *êlê* can have [-human] reference, as illustrated in (104).

(104)	<i>Zon da mu livlu</i>	>	<i>Zon da mu êlê</i>
	Zon give 1SG book		Zon give 1SG 3SG
	‘Zon gave me the book’	>	‘Zon gave me it.’

Tables 9 and 10 also suggest that changes in the pronominal paradigm apply to individual pronouns and not to paradigms as a whole (e.g. Fuß Ms.). Reduced 2sg *ô*, for instance, has unique properties compared with the other pronominals and even strong 3sg *êlê* and 1sg *ami* diverge with respect to a number of properties.

Following, for instance, Déprez (1994) for Haitian, I have assumed that at this stage Santome does not exhibit syntactic clitics. However, ultimately weak pronouns that are phonological clitics may evolve into syntactic clitics. At this point, the best candidate for this type of change is unquestionably 1sg *n*, which was shown to cliticize phonologically to the right and occurs almost exclusively in weak environments. The only property that arguably still distinguishes this pronoun from being a syntactic clitic is the intervening inclusive adverb *ten* or *tembeten*. In fact, it is difficult to find a syntactic or phonological explanation for its occurrence in this position. In a perspective of language change, it might be the case that this placement of *n* is no more than a residual case that indicates that weak pronouns at some stage of the language were able to occur in strong contexts.

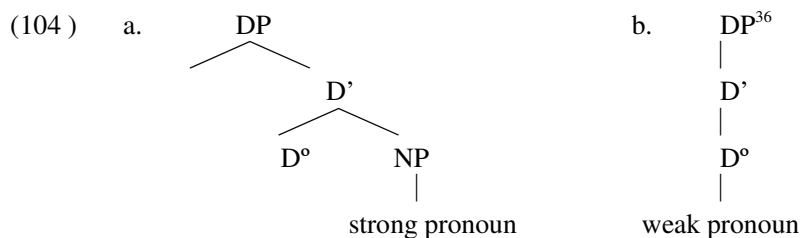
2.8.1. The internal structure of pronouns

The clitic-strong pronoun tradition initiated by Kayne (1975) gained new momentum in the middle eighties and nineties (e.g. Holmberg 1986, Vikner 1990), especially in the work by Cardinaletti & Starke (1994, 1995, 1999). Cardinaletti & Starke argue that Kayne’s syntactic and phonological model can be elegantly restated as a syntactic tripartition of the pronominal system. Therefore, a split in clitic-weak-strong pronouns is proposed, where each level implies a different degree of syntactic defectiveness.

Cardinaletti & Starke (1995: 40) claim that “(...) whenever (syntactic) clitics differ from strong pronouns with respect to some non-phonological property, the “PF clitics” pattern with syntactic clitics rather than with syntactic strong pronouns w.r.t. that (non-phonological) property.” This underlies their claim that Kayne’s PF-clitics should be relabeled as weak pronouns.

In Cardinaletti & Starke’s (1999) analysis, strong pronouns are fully projected DPs, weak pronouns lack DP-internal structure, more specifically CP, and are D°s. Based on empirical evidence, I have argued that weak forms in Santome, such as 1sg *n* and 3sg *ê*, lack support for a treatment as syntactic clitics, despite their significant degree of syntactic deficiency. Moreover, the data from Santome turn out to be relevant for the tripartite syntactic classification and in some other respects as well. A full reduced form like *am* exhibits minor syntactic deficiency. But should it be treated as a weak pronoun? If it were treated as such, it would have to be grouped together with the much weaker pronouns or, alternatively, be analyzed as weak pronouns of a distinct class (clitics) — an analysis against which I have argued above. The assumption that full reduced forms are weak pronouns also runs against an understanding that the properties of weak pronouns are much closer to those of clitics than to those of strong pronouns.

Having argued against syntactic cliticization, something needs to be said about the deficiency of pronominal forms. I consider the forms that pattern as strong forms or as both strong and weak forms full-fledged DPs. This is uncontroversial. However, from the properties discussed in section 2.4 it follows that weak 1sg *n* and weak 3sg *ê* exhibit the properties typically associated with syntactic clitics (cf. Kayne 1975). In the spirit of Laenzlinger (1998), I propose that strong and weak pronouns essentially differ with respect to their internal structure. This is conceptualized by the following representations:

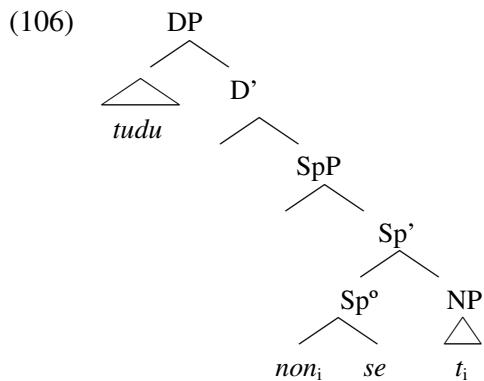


³⁶ In Bare Phrase Structure (Chomsky 1995), intermediate D' is dispensed with.

The structure for clitic pronouns proposed by Laenzlinger is the same as that for weak ones; but weak pronouns are claimed to exhibit weak phi-features, whereas clitics exhibit strong phi-features. Despite the general agreement on the fact that weak and clitic pronouns are to some extent deficient, this proposal differs from Cardinaletti & Starke's (1999) claim that weak pronouns lack a CP layer.

I assume that both weak and strong pronouns in Santome are base-generated DPs in [Spec,VP], but differ with respect to their internal structure (see the trees above). Weak pronouns are D°s that do not select a complement, whereas strong pronouns are complements of D°. I have extensively tested and discussed the difference between weak and strong pronouns, especially for the case of 1sg and 3sg (section 2.4). The test that examined the pronoun's internal structure was provided in section 2.4.8, where *ami* could be modified by specific marker *se*. This possibility applies to all strong pronouns except for *êlê*. In addition, plural pronouns can also be modified by quantificational elements, such as *tudu* 'all'. The example in (105) and the corresponding tree in (106) put the structure in (104a) into practice:

- (105) *Tudu non se.*
 all 1SG SP
 'All of us.'



Here I follow the proposal for the structure of the DP in Santome presented by Alexandre & Hagemijer (in press). The pronoun is generated as the head of the NP, which is moved and adjoins to the clitic specific marker that heads the SpP (Specific Phrase), which is a required step considering the functional structure of the DP in

Santome. The quantifier is generated as [Spec,DP]. Instead of assigning any particular syntactic deficiency to weak pronouns, such as the lack of a CP layer, I assume, following for instance Aboh (2004), that deficiency is the common ground but that languages differ with respect to the internal structure of pronouns and that even language-internally pronouns may exhibit different degrees of deficiency with a syntactic counterpart.

Having established that subject pronouns in Santome are divided into a weak and a strong class, it follows that they are XPs for syntax. This line of reasoning follows that of Cardinaletti & Starke's proposal but differs significantly with as to how the internal structure of pronouns is analyzed.

2.8.2. Summary

In the previous sections I have shown that earlier accounts of Santome's pronominal system, although relatively accurate, are incomplete. In addition to the weak-strong bias that had already been observed by most of the authors with respect to 1sg and 3sg, extensive syntactic testing resulted in system that is considerably more complex than previous descriptions suggest. I further claimed that in Santome cliticization phenomena in the pronominal paradigms are best subsumed under a phonological and not a syntactic analysis. This claim was supported by a number of arguments.

First, pronouns occupy the same positions as full-fledged DPs, i.e. there is no syntactic movement. Second, subject pronouns do not behave like agreement markers, especially because they can be separated from the verb by a diverse range of hosts, including adverb *ten* 'also'. Third, I presented evidence that weak 3sg is able to occur as the object of prepositions and in strong environments if it can cliticize to the left. Fourth, the data clearly show that phonological cliticization is a widespread feature in Santome, in both the weak and the strong paradigm. The fact that the directionality of cliticization is in most cases to the left argues against rightward syntactic (and phonological) cliticization of an agreement marker to T. The absence of a syntactic cliticization is fully consistent with the isolating character of Santome. These conclusions *per se* argue decisively against analyzing Santome as a null subject language.

2.9. Santome and the null subject parameter

From a typological point of view, syntactic clitics and the correlate that there is argumental *pro*-drop are rather unexpected features of creole languages, especially if one compares them to isolating and topic-oriented languages. Yet, DeGraff's (1993) analysis of weak pronouns as agreement heads has also been adopted for at least the following creole languages, namely Saramaccan (Veenstra 1996)³⁷, Capeverdean (Baptista 2002) and Papiamentu (Kouwenberg 2006, Veenstra forthc.). For Mauritian Creole, it has been argued that null subjects are *pro* when there is existential quantification but not when there is universal quantification (Syea 1993). This would then be a case of restricted *pro*-drop.

Several authors have rejected this type of analysis for particular languages. Déprez (1994) argues that Haitian exhibits nothing but phonological clitics and the argumentation against syntactic clitics in this language has been reinforced by Cadely (1994) and Roberts (1999). Adone (1994) claims that null subjects in Mauritian Creole are discourse-bound variables and not *pro*. Pratas (2004) argues, against Baptista (2002), that Capeverdean has weak XP pronouns that sit in [Spec,TP] and that AgrP does not even project in this language. As the foregoing makes clear, there is considerable controversy in this domain.

Taking into account previous work, I will discuss the implications of such an analysis for Santome and show that this language cannot be subsumed under the null subject hypothesis.

2.9.1. The theory of *pro*-drop

The theory of *pro*-drop has become a serious endeavor since Chomsky (1981) and has received attention in many subsequent studies. The main concern became the conditions under which *pro* can be identified and licensed, since it was readily understood that these conditions were not uniform across languages (cf. Jaeggli & Safir, 1989). The presence of verbal morphology is one such condition, but this need not be standard rich person/number agreement, as pointed out by for instance McCloskey & Hale (1984) for Irish.

Languages like Chinese, on the other hand, have very poor verbal morphology, but do exhibit *pro*-drop of subjects and objects to a significant extent, which has been

³⁷ The syntactic analysis has also been pursued for Saramaccan's main substrate language, Fongbe (Aboh 2004).

assigned to the discourse-orientation of this language (e.g. Huang, 1984, 1989). Even English, considered a typical non-*pro*-drop language, has been claimed to be a Residual Topic-Drop Language, because of a number of specific circumstances where pronouns can actually be dropped (Hyams, 1994). The results of cross-linguistic research seem to be that virtually all languages may exhibit some degree of *pro*-drop, but the conditions underlying this property are very varied.

The abovementioned creole languages arguably exhibit *pro*-drop of argumental and non-argumental subjects, despite the aforementioned absence of features like topic prominence or (rich) verbal morphology. Therefore, the source of their alleged *pro*-drop status had to be found in some other property.

Rizzi (1986) argued that certain Italian dialects exhibit clitic pronouns that should be considered agreement markers adjoined to Inflection (Tense), licensing a *pro* in subject position. This line of research has had several follow-ups with new data, especially drawn from Italian dialects (cf. Brandi & Cordin, 1989; Poletto, 2000, etc.). Little *pro* can thus be licensed in the specifier position of the head to which the syntactic clitic adjoins, say TP (or AgrP), as seen in (107):

(107) [_{TP} *pro* [_T [_T cl. [_{VP}]]]]

Yet in these dialects, the importance of rich verbal morphology plays a crucial role, a fact that is especially clear in dialects where the subject clitic takes on an invariable form.

In the case of Haitian Creole, DeGraff (1993) expands on Rizzi's line of argumentation and argues that subject pronouns are agreement markers that spell-out number and agreement features of TP (INFL in the original proposal).

Having argued against this syntactic analysis because the empirical evidence from Santome shows otherwise, I will briefly discuss whether this language exhibits cases of *pro*-drop.

2.9.2. Against *pro*-drop in Santome

In Santome, as in most Creole languages, arguments cannot be dropped in main and embedded clauses, including in the case of co-reference between the matrix and the embedded subject. Mandarin, for instance, allows *pro*-drop through co-reference with the matrix topic.

- (108) **(Ê) kopla pixi.*
 ‘He bought fish.’
- (109) *Zon kuji kuma *(ê) kopla pixi.*
 Zon answer that 3SG buy fish
 ‘Zon answered that he bought fish.’

Note also that Santome does not allow postverbal subjects linked to expletive or null subjects, a well-attested property of many *pro*-drop languages. This applies to any verb type, here illustrated with unaccusative (110) and unergative verbs (111).

- (110) a. *Ûa ngê kyê.*
 3SG person fall
 ‘Somebody fell.’
 b. **Ê kyê ãa ngê.*
- (111) a. *Ûa ngê kôlê.*
 a person run
 ‘Somebody ran.’
 b. **Ê kôlê ãa ngê.*

Santome further lacks any type of inversion, even though inversion does not necessarily correlate with the existence of null-subjects. There are, however, a few special cases discussed in more detail in the next sections.

2.9.2.1 Existential constructions

The first such case is that of existential *a/avia* ‘there is/there was/were’, clearly calqued on Portuguese existential verb forms *há/havia* ‘there is/there was/were’, from the verb *haver* ‘to be, to exist’.

- (112) *A tela ku ka tê vintxi mwala.*
 are country REL ASP have twenty woman
 ‘There are countries that have twenty women.’
- (113) *Avia ãa sungê. Nomi sun sa Velinhu.*
 was a man name POS be Velinhu

‘There was a man. His name was Velinhu.’

These two forms have the following properties: they exhibit a temporal distinction (past vs. present)³⁸, they select for an internal argument, they do not and cannot exhibit an overt expletive co-referential with the internal argument, and they are defective in the sense that they cannot be preceded by negation and TMA-markers. Despite the Case-marking property of *a* and *avia* these lexical items clearly differ from standard verbs in all the other respects above. Note that these items are particularly common in the introductory sentence of folk tales, as in (113). For all these reasons, I consider these forms grammaticalized items calqued on Portuguese. As such, they do not constitute a counterargument to the claim that Santome does not exhibit null subjects.³⁹ That *a* and *avia* cannot be subsumed under any regular paradigm follows from the properties of other existential constructions with *tê* ‘to have, to exist’, *sa* ‘to be’ and *sen* ‘to exist’, exemplified respectively in (114-116):

(114) *Ê tê ome ku sêbê kuji kume bwa so.*
 3SG exist man REL know cook food good very
 ‘There are men that know how to cook very well.’

(115) *Ê sa dôsu tan.*
 3SG be two only
 ‘There are only two.’

(116) *Ngê sen ni Putuga ku ka dumu uva ku ope.*
 People exist in Portugal REL ASP smash grape with foot
 ‘There are people in Portugal that smash grapes with their feet.’

³⁸ This morphologically encoded temporal distinction is also found in a few exceptional cases where stative verbs borrowed the lexical opposition from Portuguese, for instance *sa* ‘is’/ *tava* ‘was’ (both probably derived from Ptg. *estar* ‘to be’) ‘was’ and *kunda* ‘think’/ *kundava* ‘thought’ (from Old Portuguese *cuidar* ‘to think’). In some cases, the lexical opposition move semantically away from the etymon, e.g. *pô* ‘can’/ *pôdja* ‘should’ (from Portuguese *pode/podia* ‘can/could’). In all these cases it can be shown that the verbal properties were preserved.

³⁹ Note that calquing is also attested through the survival of subjunctive forms in fixed chunks.

- (i) *Aja vida ku sawôji.*
 exist-SUBJ life and health
 ‘Let there be life and health.’
- (ii) *Seja fêta sa vontê Dêsu.*
 be-SUBJ done be will God
 ‘Let the will of God be done.’

Unlike *a* and *avia*, these existential verbs are full-fledged verbs that can be negated, preceded by TMA-markers, exhibit an overt subject, etc. The construction with existential verb *sen* always requires a referential subject, and therefore differs from constructions in (114-115), where an overt expletive (*ê*) is required. Note that in these two constructions the postverbal DP can not be moved to the subject position and replace the expletive. These constructions are similar to English expletive constructions with ‘to be’. Hence, it can be concluded that existential constructions in Santome cannot be invoked as evidence in support of null subjects.

2.9.2.2. Unaccusative *xiga* ‘to arrive’

In section 2.9.2 I showed that Santome unaccusative verbs do not exhibit postverbal subjects. This behavior thus runs contrary to what can be observed for most Romance languages. However, there is an apparent exception to this rule with the verb *xiga* ‘to arrive’. Unlike *a/avia* in the previous section, *xiga* is not defective in the sense that it exhibits the properties of a full-fledged verb, which follows from the possibility of taking subjects, TMA-markers and negation. Nevertheless, *xiga* is special in the sense that in its intransitive use it is able to take an expletive or a null expletive subject co-indexed with a post-verbal temporal expression,⁴⁰ which is shown in examples (117) and (118) respectively.⁴¹

(117) *So xiga ja ku sun ska fe kwa se.*
 then arrive day REL man ASP do thing SP
 ‘Then the day arrived he did it.’

(118) *So ê xiga ja ku tudu ngê ska bi.*
 Then 3SG arrive day REL all people ASP come
 ‘The day arrived that everybody came.’

⁴⁰ Note that just like other motion verbs *xiga* becomes a two-place predicate with locative arguments.

(i) *Zon xiga losa.*
 ‘Zon arrived at the plantation.’

⁴¹ Note that the verb *tê* (cf. section 10.2.1) also allows for postverbal temporal expressions to form adverbial expressions:

(i) *Maji ê na tê trêxi dja fa, sode ten mal’e plôvya zôgô.*
 but 3SG NEG have three day NEG soldier also tie-3SG because game
 ‘But less than three days ago, the soldiers also tied him up because of the game.’

Null expletives are also allowed when the verb is preceded by the aspect marker *ka*, as in (119). The use of tense marker *tava* in (120) in the same construction, however, yields an ungrammatical sentence without a filled subject position in (120a). If the subject position is filled with 3sg, *ê* is interpreted as a referential subject, in (120b). In fact, arguments other than temporal expressions obligatorily occur in the preverbal position.

(119) *Ka xiga nôtxi.*

ASP arrive noite

‘The night falls.’

(120) a. **Tava xiga nôtxi.*⁴²

3SG TNS arrive noite

b. #*Ê tava xiga nôtxi*

(‘He had arrived at night’)

I did not find any cases with the verb *xiga* where the temporal expression was promoted to the preverbal position. Note that temporal expressions typically occupy the subject position of the clause, which follows from the grammaticality of (121a-b) and the ungrammaticality of (122a-b).

(121) a. *Ûa dja ka be, ùa dja ka be ku nantan ka bila-bila fa.*

one day ASP go one day ASP go REL never ASP turn-turn NEG

‘Every day that goes by will never come back.’

b. *Plaman ka bili, nôtxi ka kubli.*

Morning ASP open night ASP cover

‘The morning breaks, the night falls.’

(122) a. **Ê ka be ùa dja, ê ka be ùa dja ku nantan ka bila-bila fa.*

b. **Ê ka bili plaman, ê ka kubli nôtxi.*

In modern European Portuguese, when *chegar* ‘to arrive’ (the etymon of Santome *xiga*) involves physical arriving, the preverbal and postverbal positions of the subject are

⁴² This sentence can only read as ‘He had arrived at night.’

typically regulated by information structure. New information prefers the postverbal position, old information the preverbal position. Temporal expressions with subject interpretation, on the other hand, show a preference for the postverbal position and are less bound to information structure. Given these facts, it is reasonable to argue that calquing on Portuguese is responsible for the observed pattern. All in all, the exceptional behavior of *xiga* cannot be considered an argument in support of the claim that Santome exhibit null expletives.

2.9.2.3. Copular verb *fika*

Another special case of arguable non-argumental *pro*-drop is the copula verb *fika* ‘to remain’. The subject of this verb can stay in post-verbal position through coindexation with a null expletive or an overt expletive, in (123) and (124) respectively. Contrary to the case of *xiga*, the postverbal subject can be raised to the preverbal subject position, as shown in (125).⁴³

(123) *Fika dôsu ome tan.*
 remain two man just
 ‘There remained just two men.’

(124) *Ê fika dôsu ome tan.*
 ‘There remained just two men.’

(125) *Dôsu ome tan fika.*
 ‘Just two men remained.’

⁴³ The verb *fika* has two meanings, namely intransitive ‘to remain’ and transitive ‘to leave (behind)’. In both cases, the subject has to be realized. In a special case, the subject has to be omitted.

(i) *Likêza ku n tê, n ga môlê fika da bô.*
 resources that 1SG have 1SG ASP die leave give 2SG
 ‘The resources that I have, I will leave to you upon my death.’

In this sentence *môlê fika* behaves like a lexicalized serial verb construction. In enumerations, repeated subjects can be dropped but the absence of a discourse pause and the impossibility to independently negate *fika da bô* show that we are not dealing with an enumeration, such as the following:

(ii) *Ûa môlê, fika trêxi tan.*
 One die remain three only
 ‘One (child) died, only three remained.’

Previous discourse identifies the subject of *fika*. Note also that a clear discourse break is present.

As in the case of *xiga*, *fika* exhibits the full range of functional projections associated with verbs. This follows for instance from the fact that negation, which sits higher than TMA-markers, can be projected with an empty subject position:

- (126) *Na fika nê ãa planta sungê munjadu fa.*
 NEG remain not one plant man stand-PP NEG
 ‘There remained no plant of the man in upright position.’

In sum, I assume that *fika* exhibits the properties of a raising verb, as in the superstrate, and should be considered one of the exceptional cases where null expletives can be used in Santome, since other raising verbs, for instance certain modal and aspectual verbs, do not license null expletives.

2.9.2.4. A raising verb?

Santome exhibits the item *palêsê* ‘apparently, it seems’, which is etymologically derived from the Portuguese raising verb *parecer* ‘to seem’:

- (127) *Palêsê êlê ku mosu dê ka bi n’ãa vapô.*
 Seem 3SG with boy POS ASP go in-a ship
 ‘Apparently he and his boy will come in a ship.’

But there is hardly any similarity between this item in both languages. In (128), it is shown that the embedded subject *ê* cannot be promoted to the clause-initial position. Moreover, *palêsê* cannot receive any TMA-marking, in (129a), nor be negated, in (129b). TMA-marking and negation always have to occur in the clause that follows *palêsê*.

- (128) a. *Palêsê ê ska dwêntxi.*
 seem 3SG ASP ill
 ‘It seems he is ill.’
 b. **Ê palêsê ska dwêntxi.*
 (He seems to be ill)
- (129) a. **Taval*ka palêsê ê ...*
 TNS/ASP seem 3SG

- b. *Na palêsê ... fa
 NEG seem ... NEG

The embedding relation between the two clauses in the following example also shows that *palêsê* does not select a complement as in Portuguese.

- (130) *Palêsê pa a na fut'e djêlu, ê fisa poto.*
 Apparently for IMP NEG steal-3SG money 3SG close door
 'Apparently, so they wouldn't steal money from him, he closed the door.'

In fact, *palêsê* occurs on top of a purpose clause, which follows for instance from the fact that there is no final negation marker in this construction (section 4.2.2). In sum, *palêsê* cannot be considered a raising verb like its Portuguese counterpart and exhibits the behavior of an evaluative modal adverbial.

2.9.2.5. Discourse-bound topic drop

The question-answer structure in (131) suggests that Santome exhibits topic-drop in specific question-answer environments, contrasting with 'out-of-the-blue' sentences, where subjects cannot be dropped, as illustrated in (132).

- (131) Q: *Kê kwa ku ê/inen ka fe?*
 what thing KU 3SG/3PL ASP do
 'What does he/they do?'
 A. *Ka kume ka bêbê.*
 ASP eat ASP drink
 'He/They eat(s) and drink(s).' / 'Eating and drinking.'
- (132) #*Ka kume ka bêbê.*
 ASP eat ASP drink
 *X eat(s) and drink(s).
 ('Eat and drink!')

Note that (131A.) can be used as an imperative. According to the informants consulted, these null subjects are thus only licensed under two conditions: (i) when they recover a 3sg or 3pl pronoun, and (ii) in the presence of aspect marker *ka*. In Chapter 3 it will be

shown that *ka* has the status of a bound morpheme, which arguably explains why it can occur in this specific construction, unlike the other TMA-markers.⁴⁴ The answers to the questions in the following examples are therefore considered ungrammatical:

(133) Q: *Kê kwa ku ê ta fe?*
 what thing KU 3SG TNS do
 ‘What had he done?’

A: **Ta kume ta bêbê*
 TNS eat TNS drink

(134) Q: *Kê kwa ku bô ka fe?*
 what thing KU 2SG ASP do
 ‘What do you do?’

A: **Ka kume ka bêbê.*
 (OK: *N ga kume n ga bêbê.* ‘I eat and drink.’)

This is also one of the contexts where Mauritian allows null subjects with definite reference, which is assigned to the specific discourse context in which the answer can only refer to the subject in question.

(135) Q: *ki Pyer pe fer?* (Mauritian; Syea, 1993: 93)
 what Peter ASP do
 ‘What is Peter doing?’

A: *Pe petir labutik*
 ASP paint shop
 ‘He is painting the shop.’

However, as follows from Syea (1993) and Adone (1994), in Mauritian subject drop is not limited to 3rd person, but covers the whole pronominal paradigm and the presence of TMA-material is required to drop subjects. In Adone’s analysis, null subjects are variables linked to a discourse topic. Since the licensing of empty subjects in Santome is far more restricted than in Mauritian, i.e. it only occurs with 3sg/pl and in the presence of one specific TMA-marker, I will pursue neither a *pro*-drop nor a variable analysis.

⁴⁴ In section 9.2.2 it was shown that a similar constraint applied to *xiga*.

Instead, I propose that the *ka+V-ka+V* construction should be subsumed under Control Theory, whereby a discourse-bound *PRO* has to be identified as a participant, but not the speaker nor the addressee. Despite the specific referentiality, I propose that the Santome construction is similar to the following English and Portuguese construction:

(136) [PRO] *Eating and [PRO] drinking.*

(137) [PRO] *A comer e [PRO] a beber.* (European Portuguese)

‘Eating and drinking.’

‘Eat and drink!’

It is perhaps no coincidence that the Portuguese construction can also have an imperative meaning, similarly to (132) above. In sum, I do not consider the structure investigated in this section an instance of referential *pro*-drop.

2.9.2.6. Enumeration & coordination

Ferraz (1979: 65) notes that “the [3sg] pronoun may optionally be deleted when the subject is repeated in time sequence in a co-ordinate clause.” These are cases of deletion under identification and, differently from the findings in the sections above, this feature targets argumental subjects. It should also be noted that my data show that this process is not limited to 3sg, as illustrated in (138-140).

(138) [*Inen*]_i *ka pya sun, [-]_i ka li.*

3PL ASP look at man ASP laugh

‘They looked at the man and laughed.’

(139) [*Sun*]_i *ka tlabá sun, [-]_i tê kwa sun, [-]_i bêbê ãa vinpema sun.*

man ASP work PSR have thing POS drink a palmwine POS

‘He does his job, has his things, drinks his palmwine.’

(140) [*N*]_i *ga subli ê plaman, [-]_i bila subli taji...*

1SG ASP climb 3SG morning turn climb afternoon

‘I climb it [palmtree] in the morning and then again in the afternoon...’

Since Santome lacks an overt coordination conjunction for coordinations at VP level or higher, these are typically asyndetic structures.⁴⁵ The examples show that non-realized subjects can be licensed by any pronoun, and that coordination may take place at different levels, for instance VP in (139-140) or AspP in (138).

The discussion on coordination structures has played a relevant role in the literature on subject clitics. Rizzi (1986) argues that one of the properties of subject clitics derives from the fact that they have to be repeated in each conjunct. This leads him to argue that in Trentino (141b) *egli* in (141a) is not a clitic, whereas *la* in (141b) is.

- (141) a. *Egli canta e [-]_i balla benissimo.* (Standard Italian)
 ‘He sings and dances very well.’
 b. *La canta e *(la) bala.* (Trentino)
 ‘She sings and dances.’

This argument has also been used in creole studies. Baptista (2002: 260-61) claims that an identical restriction applies to Capeverdean in support of the claim that this language has subject clitics. (Capeverdean; Baptista 2002: 261)

- (142) a. *João bebe se vinhu i [-]_i bai se kaminhu.*
 João drink his wine and go his way
 ‘João drank his wine and went his way.’
 b. *E bebe se vinhu i *(e) bai se kaminhu.*
 João drink his wine and (he) go his way
 ‘João drank his wine and went his way.’ (Capeverdean; Baptista 2002: 261)

As in the Italian examples above, coordination by XP (*João*), in 142a), licenses a null subject in the second conjunct, whereas the alleged clitic (*e*), in (142b), does not.

Baptista further considers that enumerations in Capeverdean constitute an additional argument for a *pro*-drop analysis because subject repetition is not required:

⁴⁵ Sometimes the Portuguese conjunction *e* ‘and’ is employed in coordinate structures, but this is clearly a calque on Portuguese. There is no predictable pattern associated with the use of this conjunction.

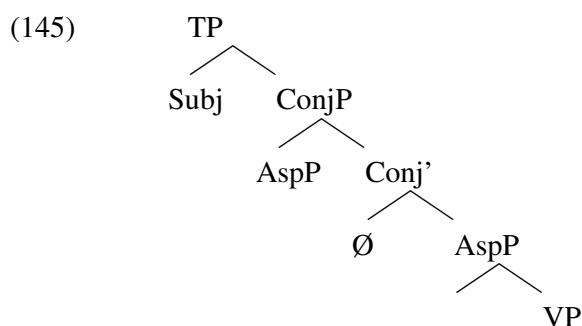
(143) *N bai nha kaza dja, N bai pega na vivensia pa [-]konta na nha kaza. [-] Panha lenha na montadu, [-] bende ... [-] Bende kel fixinhu di lenha, [-] ba trabadja djenti, [-] ganha kel dinhirinhu, bem kumpra kel kafizinhu.*

‘I went to my own house then, and went to seek a livelihood relying on my home. I would take the wood in the grove and would sell it. I would sell that little piece of wood, I would go to work over people’s houses, I would earn a small sum, I would buy a little coffee.’ (Capeverdian, Baptista 2002: 259)

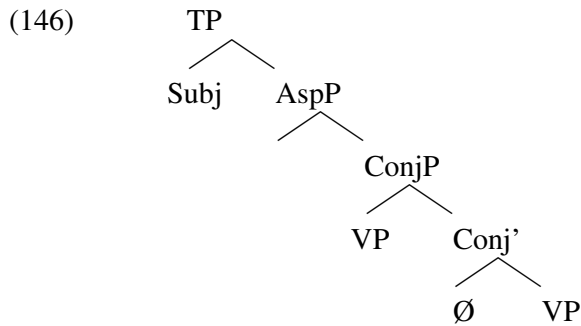
According to the classification of Baptista (2002), Capeverdean 1sg pronoun *n*, like 3sg *e* in (142b), is a syntactic clitic.⁴⁶ Since enumerations are asyndetic coordinations and usually subsumed under the same type of analysis as asyndetic coordination (e.g. Matos, 2003), the Capeverdean data should be considered contradictory. Contrary to fact, the prediction is that there should be repeated subject clitics in enumerations as well. In fact, a typical non-null subject language such as English also exhibits enumerations lacking a subject, as can be seen from the following complex event.

(144) *John opened the fridge, took the beer, cut the cheese, went to the living and settled on the couch.*

It follows that enumerations do not form true cases of *pro*-drop but reflect the structure of coordinated clauses. Anticipating several facts about clause structure in Santome that will be discussed in Chapter 2, I assume that an example like (138), involving coordination at the aspectual level, has the abbreviated representation in (145), whereas cases of VP-coordination, such as (139-140), can be represented as in (146).



⁴⁶ Cf. also Pratas (2002).



In these structures, the subject sits arguably in [Spec,TP] and triggers an across-the-board effect with respect to the conjuncts.

The purpose of this section was to show that the structures above are not a reliable test to prove that a given language has syntactic clitics. While it is fairly uncontroversial that Capeverdean exhibits non-argumental *pro*-drop (cf. Baptista 2002: 254-5), coordination/enumeration structures cannot be used as an argument for argumental *pro*-drop.

2.9.2.7. Serial verbs

I will conclude this section with a short note on serial verb constructions. Serial verbs are very productive in Santome (Hagemeijer 2000, 2005b), and are well known for the fact that they allow only one (initial) subject in VP₁, typically (though not always) the understood subject of VP₂.

Byrne (1985) put forth the hypothesis that VP₂ in Saramaccan takes a *pro* as its subject. Borer (1989) briefly expands upon this analysis and subsumes it under the wider claim that languages may exhibit anaphoric agreement. According to this hypothesis, the lower AgrP (assuming that AgrP is indeed projected) in Saramaccan would move to COMP where it is bound by the matrix subject and receives its features. However, the need for an embedded CP position in Saramaccan's SVCs, and arguably in all SVCs cross-linguistically, does not receive any empirical motivation.

Veenstra (1996) provides an exhaustive description and analysis of SVCs in Saramaccan and concludes that the lower event is an adjoined AspP with a PRO coindexed with the matrix subject. Since the data of Saramaccan SVCs match the Santome findings to a significant extent, in Hagemeijer (2000) I retained the essence of Veenstra's (1996) analysis. Consequently, the lower subject position in Santome SVCs is not a counterargument to the non-*pro*-drop status I have been arguing for.

2.9.3. Interpreting the *pro*-drop data

The data survey in the preceding sections showed that true *pro*-drop is a highly restricted phenomenon in Santome and is crucially limited to non-arguments. Other languages that are claimed to exhibit non-argumental *pro*-drop are, for example, German or Icelandic. It is also generally accepted that non-argumental *pro*-drop does not necessarily entail argumental *pro*-drop (cf. Jaeggli & Safir, 1989).

For the present purpose, it is, once again, revealing to look at Haitian, not only because subjects in this language have received the careful scrutiny of several scholars (e.g. DeGraff 1993, Déprez 1994, Cadely 1994, Lefebvre 1998), but also because Haitian shares an extensive number of typological similarities with Santome, such as preverbal marking, serialization, poor inflectional morphology, negative concord, non-topic-orientation, etc. Furthermore, both languages are plantation creoles, and are derived from a Romance lexifier whose main substrates share numerous typological features.

If we look at the data in DeGraff (1993) and Déprez (1994), it follows that Haitian allows for null subjects with raising verbs like *rète* ‘to remain’ and *sanble* ‘to appear’ and *gen* ‘to be’. Adjectival predicates and meteorological verbs without arguments require an obligatory expletive in this language.

Note that Santome obligatorily uses overt expletive subjects (*ê*, *kwa*) with adjectival predicates.

- (147) *Ê na sa bwa fa ô.*
3SG NEG be good NEG EMPH
‘It’s not good!’

When the adjectival predicate selects a clause, this clause will typically be promoted to the subject position.

- (148) *Tlaba ku mafe na bwa fa.*
work with mistrust NEG good NEG
‘It is not good to work mistrustingly.’
- (149) **Ê na bwa tlabu ku mafe fa.*
3SG NEG good work with mistrust NEG

Furthermore, irrespectively of their argumental status, meteorological verbs cannot take null subjects in Santome:

(150) *(Ê) *ka sôbê muntu fan!*
 3SG ASP rain a-lot EMPH
 ‘It rains a lot!’

(151) *(Ê) *ka venta.*
 3SG ASP be windy
 ‘It is windy.’

These cases thus set Haitian and Santome apart. Despite this difference, the case of meteorological verbs in Haitian is worth exploring further because it is crucial to the analysis of Déprez (1994: 3), which will be discussed below. Note the following contrast.

(152) a. (*Li*) *fê lapli.* (Haitian; Déprez 1994: 3)
 3SG make rain
 ‘It is raining.’
 b. *(*Li*) *vante.* (Ibidem)
 ‘It is windy.’

As pointed out by Déprez (1994), the non-argumental null subjects in Haitian should not be treated on a par with non-argumental subject drop of the German type, where this phenomenon has been analyzed as the result of V to C movement, since Haitian does not exhibit this type of movement. In order to explain the contrast in (152a-b) above, Déprez adopts the analysis proposed by Borer (1986) that non-argumental subject positions do not have to be projected in syntax, extending to the point that coindexation between INFL and the postverbal argument suffices to satisfy the EPP. Hence, in the presence of an argument, as in (152a), the expletive does not need to be projected, whereas in (152b) no coindexation can take place and thus requires projection of an overt expletive.

The question is now whether this constraint on non-argumental *pro*-drop can be upheld for Santome. The examples I have discussed in the previous sections are, in fact, compatible with the hypothesis that expletives are overt when there is no argument

available to fill in the subject position. This could be seen with meteorological predicates and adjectival predicates. Note also that the sharp contrast found with Haitian meteorological verbs does not carry over to Santome, where the type shown in (152a) is inexistent. It also predicted that arguments in postverbal position, i.e. raising verbs and the *xiga*-cases, do not require overt expletives.⁴⁷

2.9.4. Summary

In this section I have reviewed potential cases of *pro*-drop in Santome and concluded that there are only sporadic cases where non-argumental *pro*-drop occurs. In the absence of arguments in support of syntactic clitics as agreement markers, the general conclusion is that Santome is a well-behaved non-*pro*-drop language.

⁴⁷ Apparently, some uses of *sa* ‘to be’ occur with null subjects. However, it turns out that this is only possible when it functions as a highlighter in cleft constructions (cf. (i)). I consider this an instance of grammaticalization, since expletives cannot precede *sa* in these structures. When *sa* behaves as a copula verb, the subject position has to be filled by an overt expletive (cf. (ii)).

- (i) (* \hat{E}) *sa vin ku solo na ka lentla fa.*
 3SG be wine that sun NEG ASP enter NEG
 ‘It is wine where the sun doesn’t get into.’
- (ii) *(\hat{E}) *sa vede.*
 3SG be truth
 ‘It is true.’

3. THE EXTENDED VERB PHRASE

3.1. Introduction

The first part of this chapter focuses on the position of the verb in the clause. It has been argued for most creole languages that the verb remains *in situ* and that poor inflectional morphology is the trigger for the lack of verb movement. I will draw on several standard tests as well as on language-internal evidence in order to determine how Santome behaves with respect to this property.

The second and largest part of this chapter will be concerned with the TMA-system of Santome. Despite the fact that TMA-systems have always been a popular topic in creole studies, the previous literature on Santome has left many aspects of its TMA-system untouched.

This chapter has the following outline: Section 3.2 discusses verb movement. In addition to standard tests such as adverb placement and quantifier float, evidence is drawn from double object constructions (DOCs) and the syntactic particularities of the allomorphs *be* and *be*, both meaning ‘to go’. Section 3.3 focuses on the TMA-system. First, I will briefly review earlier proposals, followed by a description of the core temporal-aspectual material and the combinations thereof. Section 3.3.3 provides an in-depth account of stativity and in section 3.3.4 it is summarily shown that Santome is a relative tense language. In the light of the descriptive findings, section 3.3.5 focuses on the syntax of tense and aspect in this creole.

3.2. The verb phrase and verb movement

In this section I will investigate aspects of the Verb Phrase (VP) with special focus on verb movement. In particular, I will use several diagnostic tools to show that Santome does not exhibit any verb movement to functional projections higher up in the clause, which is in agreement with the findings for most other creole languages (cf. Roberts 1999).⁴⁸ I will show that in a rigid syntax language like Santome, standard tests for verb movement, such as adverb placement and quantifier floating, do not necessarily constitute sufficient evidence for the lack of verb movement because they do not exclude the possibility of, for instance, short movement to an aspectual node

⁴⁸ According to the proposals of Baptista (2002) and Rottet (1992), Capeverdean and Louisiana Creole, respectively, do exhibit verb movement.

immediately above VP. Therefore, I will provide a detailed discussion of a language-internal construction that sheds new and more conclusive light on (the absence of) verb movement, namely the comitative-locative shift that occurs with the allomorphs of ‘to go’, *ba* and *be*.

3.2.1. Verb movement

The relation between inflectional morphology, split IP and verb movement has been extensively treated in the literature since Pollock’s (1989) split IP proposal. Work by Roberts (1993), Holmberg & Platzack (1995), Vikner (1995) and Rohrbacher (1999) focusing essentially on Germanic languages has emphasized the relation between the presence of inflectional morphology and verb movement.⁴⁹ In short, poor inflectional morphology does not trigger overt verb movement to T°, a claim that has been carried over to creole languages (e.g. Veenstra 1996, DeGraff 1997, Roberts 1999).

A different line of analysis has been pursued by Thráinsson (1996), Bobaljik (2000) and Bobaljik & Thráinsson (1998). The essence of their proposal is that functional projections do not follow a universal pattern, i.e. languages may cross-linguistically vary with respect to what is or can be projected.⁵⁰ Their working hypothesis is grounded in the claim that rich verb morphology implies the existence of a split IP. In Bobaljik (2000), it is explicitly assumed that the presence of more than one identifiable inflectional affix on the verb stem correlates with the presence of a split IP, whereas one or no inflectional affix would typically correlate with a non-split IP, although this is not a necessary entailment. These generalizations have led Baptista (2002) to claim that Capeverdean creole is an exceptional language in the sense that it has a split IP but only one verbal affix (anteriority marker *-ba*). Baptista argues therefore that Capeverdean exhibits verb movement to T.⁵¹ That poor verb morphology does indeed not necessarily entail absence of verb movement is also evident in Afrikaans, where the Dutch verb-second phenomenon survived (e.g. Robbers 1997). Veenstra (2006a) further suggests that Papiamentu and Berbice Dutch exhibit V-movement as well, but targeting different projections, respectively Pred and Asp. Hence, he concludes that there is at least a four-way split with respect to verb movement

⁴⁹ For a comprehensive overview, I refer the reader to Koenenman (2000).

⁵⁰ This claim is of course extendable to other domains of grammar as well. Gonçalves (1999), for instance, shows that complex verbal predicates (e.g. causatives) in European Portuguese have a defective functional structure.

⁵¹ This hypothesis has been argued against by Costa & Pratas (2004) and Pratas (2004), who propose that postverbal *-ba* is lowered in morphology.

in creole languages: no movement (e.g. Saramaccan), V-to-I (e.g. Capeverdean), V-to Pred (Papiamentu) and V-to-Asp (Berbice Dutch). In the following sections I will argue that Santome belongs to the first type.

3.2.1.1. Adverb placement

A standard diagnostic for verb movement is adverb placement (cf. Pollock 1989). It will be seen in the section on TMA-marking that there are significant restrictions on the distribution of adverbs in Santome. From (1-2), it follows that typical VP-adverbs like *ndjandjan* ‘quickly, fast’ and *ben* ‘well’ can only occur in VP final position.

- (1) a. *Ê bili zanela ndjandjan.*
 3SG open window quickly
 ‘He opened the window quickly.’
 b. **Ê bili ndjandjan zanela.*
 c. **Ê ndjandjan bili zanela.*
- (2) a. *Ni wê sungê, san ka tlata mina se ben.*
 in front man she ASP treat child SP well
 ‘In front of the man, she treats the child well.’
 b. **Ni wê sungê, san ka tlata ben mina se.*
 c. **Ni wê sungê, san ka ben tlata mina se.*
 d. **Ni wê sungê, san ben ka tlata mina se.*

The fact that in no circumstances can the object be separated from the verb by an intervening adverbial is a first, strong indication that the verb does not move to a higher functional projection. The relation between the verb and the object is strictly local.

Assuming that adverbs like *ben* ‘well’ and *ndjandjan* ‘quickly’ are effectively VP-adverbs raises the tricky question as to what positions adverbs occupy in clause structure. Mainstream analyses continue to treat adverbs as adjuncts, but there is a great deal of controversy with respect to the locus of adjunction. On the one hand, Kayne (1994) proposed the Linear Correspondence Axiom (LCA) according to which languages all follow a universal specifier-head-complement order and the distinction between specifiers and adjuncts is suppressed. What matters most for the present discussion is that under Kayne’s hypothesis adverbs are always adjoined to the left. A

direct and obvious consequence of this assumption is that movements are required to derive correct word orders.

Let us consider VP-adverbs in Santome. In Kayne's (1994) framework, VP-adverbs are left-adjoined to VP, but this makes wrong predictions about word order and requires the VP to somehow move across the adverb in order to derive the correct surface order. In some languages, e.g. languages with rich inflectional morphology, this is usually an uncontroversial assumption, since movement of the verb and the object out of their basic position is theoretically motivated by feature checking and by picking up inflectional material in order to derive the correct surface word order. In a language like Santome, however, strict left-adjunction should be ruled out on empirical and theoretical grounds. On the one hand, adverb placement is highly restrictive. Adverbs cannot, for instance, be stacked in between preverbal functional TMA-heads, which leads to the conclusion that certain positions are simply opaque for adjunction or adverb placement because of particular requirements of clause structure.

Cinque's (1999) approach, on the other hand, departs from the assumption that adverbs reflect the presence of functional projections in the architecture of the clause. In his view, adverbs are the specifiers of these projections. This hypothesis not only entails a significant increase of the functional structure of a clause due to lexical items that typically do not belong to a closed-class, but also assumes that these functional projections generally lack an overt head. In Cinque's view on clause structure, the adverb *djandjan* is the specifier of a projection for celerative aspect (AspP_{celerative}). In English, the specifier of AspP_{celerative} would be *fast* and *quickly* which, according to Cinque, may occur in a higher or a lower position in the light of, for instance, the following patterns in English:

(3) *He {quickly/*fast} ran home {quickly/*fast}.*

First, it is awkward that the same functional projection should be able to sit in different positions. Second, since the functional projections proposed by Cinque project above VP, implementing this hypothesis requires cyclic movements of portions of the clause. For instance, the VP is raised to a given AspP and then this AspP is raised to another, higher AspP, and so on. For a more syntactic application of Cinque (1999), I refer the reader to Durrelman (2000), who provides a detailed account of adverbs in Jamaican

Creole. For discussion of the adverbs-as-specifier hypothesis in general, I refer the reader to Costa (2000) and Ernst (2002).

Although less restrictive, right-adjunction in the sense of Ernst (2002) significantly simplifies the derivation of adverbs within clause structure.⁵² Right-adjunction to VP dispenses with the need for verb (and/or object) movement and accounts for ordering patterns in a straightforward way together with the diagnostics for verb movement, which will be discussed in the next section. Throughout this dissertation, I will be using right-adjunction as a means to derive word order patterns.

3.2.1.2. Quantifier float

Quantifier floating is another standard test to determine whether a given language exhibits verb movement. Analyses of quantifier float come in several types. In the original proposal (Kayne 1975, Maling 1976), the quantifier is stranded from the noun with which it forms a unit, and adjoins to a maximal projection.

The influential proposal of Sportiche (1988) considers floated quantifiers to be the result of NP movement to the left. Under this type of analysis, an empty trace corresponding to the moved NP follows the quantifier and is coindexed with the NP in a standard fashion. Since the quantified DP is arguably base-generated in [Spec,VP], it immediately follows that the verb must have been raised across the quantifier in post-verbal position.

Another line of research on floating quantifiers does not follow the syntactic assumptions above, i.e. that the quantifier and the modified noun form a syntactic unit that is split (Bobaljik 1995). Rather, it is assumed that quantifiers are adjoined adverbials that are linked to the NP through an interpretation rule. Hence, this type of analysis cannot be considered a diagnostic tool for verb movement. However, it makes better predictions with respect to other quantificational elements (e.g. *few*, *every*, *no*, etc.) that typically cannot occur in the same position as canonical *all*.

However useful the quantifier floating test may be in many languages, in Santome it has to be ruled out on independent grounds and therefore does not constitute a reliable diagnostic for verb movement. Santome exhibits no bare quantifiers, i.e. quantifiers that occur without a host-DP, as illustrated by quantifiers *tudu* ‘all’, *kada*

⁵² I refer the reader to Costa (1998) for a critique of right-adjunction.

‘all’ and *kwakwali* ‘any’. Therefore, the two classic analyses of floating quantifiers outlined above are simply not feasible.

- (4) *Tudu/kada/kwakwali *(sode) ba matu.*
 All/each/any soldier go bush
 ‘All the soldiers/Each soldier/Any soldier went to the bush.’

- (5) *Sode ba (*tudu) matu (*tudu).*
 # ‘The soldiers went all to the bush-bush.’

Note further that Santome also exhibits quantifier *tudaxi* ‘all’, which additionally means ‘also, everybody’, according to the syntactic position in which it occurs, as illustrated in the following examples.

- (6) a. *Tudaxi non ka blêgê.*
 everything 1PL ASP eat
 ‘We eat everything.’
 b. *Bô ka ba paga tudaxi.*
 2SG ASP go pay everything
 ‘You will pay everything.’
- (7) *Ami tudaxi sêbê kontaji se.*
 1SG also know tale SP
 ‘I also know that tale’
- (8) *Inen ngê se tudaxi sa dôdô.*
 3PL people SP all/also be nuts
 ‘These people are all nuts.’

In examples (6a-b), *tudaxi* in subject and object position is a nominal⁵³, whereas in (7), it is an inclusive focalizer, as mentioned in section 2.4.3. When it follows a plural noun,

⁵³ *Tudaxi* is presumably a contracted form of *tudu+axi* (lit. *all+like this*). A similar relation can be found in *nadaxi* ‘nothing’ derived from *nada+axi* (lit. *nothing+like this*).

- (i) *Ami na tê nadaxi di da nansê fa.*
 1SG NEG have nothing for give 2PL NEG
 ‘I don’t have anything to give you.’

as in (8), ambiguity arises between a quantificational reading and an inclusive focus reading. However, there is no reason to believe that *tudaxi* in its quantificational use exhibits properties typically assigned to floating quantifiers. *Tudaxi* cannot precede the noun that it modifies nor can it occur in postverbal position.

(9) **Tudaxi ngê sa dôdô.*

all people be nuts

(10) **Inen ngê se sa tudaxi dôdô.*

3PL people SP be all nuts

Thus, irrespective of the theoretical standings that one adopts, the quantifier floating test has to be dismissed in the case of Santome. Thus, the impossibility of floating and the facts from adverb placement reinforce the hypothesis that verbs do not move in Santome.

3.2.1.3. Double object constructions

In this section I will focus on DOCs and especially on pronominalized direct and indirect objects. A significant number of studies have proposed that clitics are generated as functional heads above VP (e.g. Sportiche 1996). Irrespective of the label assigned to the projection hosting the clitic, the implication is that, in the case of verbal clitics, the verb moves out of its shell to adjoin to the clitic. It follows that an account along these lines cannot be consistent with the concept of V-in-situ. Hence, the discussion of DOCs can be considered as a follow-up of section 2.6.2, where I argued that Santome exhibits phonological cliticization exclusively.

As in most creole languages (cf. Bruyn, Muyken, and Verrips 1999), DOCs are quite widespread in Santome. In addition to the DOC with full-fledged verbs, in (11-12), semantically light verbs, such as *da* and *ligi* in (13-14), are also able to select two complements.⁵⁴

⁵⁴ Note that DOCs in Santome do not always grant support to a Small Clause analysis (Kayne 1984) according to which both objects are arguments and are linked through a silent element equivalent to BECOME or HAVE. In fact, there is often no predication relation between both objects, i.e. it is not always the case that the subject acts upon the IO in order to transfer the DO to the IO, since the subject of the double object verb can function as the benefactive/recipient and the IO becomes the experiencer, as for instance in example (14). For discussion of other problems with the SC analysis, I refer the reader to Pesetsky (1995: 157-63).

- (11) *Ê piji [san lenha] [kopu d'awa].*
 3SG ask lady queen glass of-water
 'He asked the queen for a glass of water.'
- (12) *Flêgadô bolo [mosu] [mindjan].*
 Massager rub boy remedy
 The massager massaged the boy with a remedy.'
- (13) *Non ka da [pema] [faka].*
 1PL ASP give palm tree knife
 'We cut an opening in the palm tree.'
- (14) *Ê ligi [palayê] [pixi].*
 3SG lift up market seller fish
 'He stole fish from the market seller.'

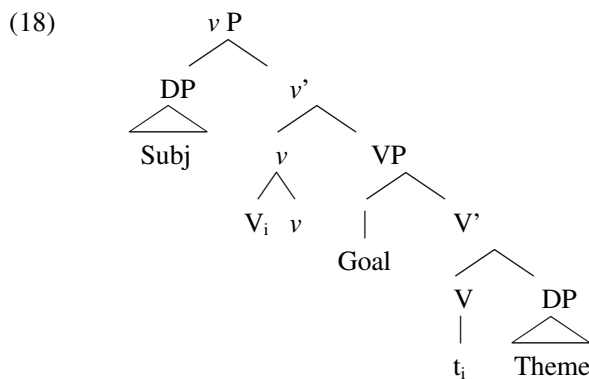
First, it should be observed that from a linear perspective the indirect object *systematically* precedes the direct object, to the extent that there is no Heavy NP Shift (Larson's 1988 Light Predicate Raising), as follows from the ungrammaticality of (15b). If anything, the heavy constituent is fronted (15c).

- (15) a. *A tlega [sun se ku ska ba fla ku Zon oze] [mina].*
 IMP give man SP REL ASP go speak with Zon today child
 'They handed the child over to the man who went to speak with Zon today.'
- b. **A tlega [mina] [sun se ku ska ba fla ku Zon oze].*
- c. *[Sun se ku ska ba fla ku Zon oze]_i, a tlega [sun]_i [mina].*
 'The man who went to speak with Zon today, they handed him over the child.'

Larson (1988) demonstrates extensively that the IO c-commands the DO. Tests with anaphors (16) and quantifier-pronoun binding (17) show that this asymmetry also holds in Santome.

- (16) a. *N musa piskadô ubwê dê me ni supe.*
 1SG show fisherman body POS self in mirror
 ‘I showed the fisherman his own body in the mirror.’
 b. **N musa ubwê dê me piskadô ni supe.*
- (17) a. *Bô da tudu tlabadô_j lôpa dê_{j/k}*.*
 2SG give every worker clothes POS
 ‘You gave every worker his clothes.’
 b. *??/*Bô da lôpa dê_{j/k} kada tlabadô_j.*

In addition to the facts from binding relations, the DO is typically considered a more intimate object of the verb, which follows, for example, from facts related to idiom chunks and the surfacing of IOs as PPs cross-linguistically.⁵⁵ As a consequence, several influential proposals (e.g. Larson 1988, Bowers 1993) have adopted an analysis whereby the IO sits in a specifier of VP, while the verb and its DO are sisters. Since for the present purpose it is not directly relevant whether the transitive structure is construed around a predicative projection (Bowers 1993), zero affixation (Pesetsky 1995), an applicative projection (Marantz 1993) or both an applicative and a transitive projection (Collins 1997), I have adopted a light verb (*v*P) shell (Chomsky 1995), in (18), where V can check off its features, yielding the correct surface order and c-command relation between IO and DO. This representation can be considered a recast of the proposal of Larson (1988) and Bowers (1993).



⁵⁵ Therefore, it has often been assumed that IOs are selected by a 'silent' preposition. Overt evidence for prepositions that become silent can be found in incorporating languages (Baker 1988).

Moreover, there is no reason to believe that we are dealing with a structure that is transformationally related to V-DP-PP (cf. Baker 1988), i.e. dative shift and preposition incorporation. Aboh (2004) uses DOCs in Gbe as an additional argument for verb movement in this language. In Gungbe, both the theme-goal and the goal-theme are grammatical, but when one of the objects is pronominalized the pronoun has to be adjacent to the verb. Under Aboh's syntactic approach, the clitic therefore moves as an X° to a higher projection, Agr° . The verb, in turn, is moved and left-adjoins to the clitic in Agr° , on its way to AspP .

In Chapter 2 it was shown that, in Santome, a phonological approach to pronouns fares better than a syntactic approach, which is confirmed by DOCs. We have seen that full objects cannot shift places and this claim also holds for pronominalized objects. In Chapter 2 it was already shown that pronominal Themes that are not adjacent to the verb require the occurrence of a strong pronoun, *êlê*. Consider the following examples in which it is shown that the sequences IO-DO in (19), IO_{clitic}-DO (in 20), IO-DO_{clitic} (in 21) and IO_{clitic}-DO_{clitic} in (22) cannot be switched with each other.

- (19) a. *Ê da Zon livlu se.*
 3SG give Zon book SP
 ‘He gave Zon the book.’
 b. **Ê da livlu se Zon.*
- (20) a. *Ê d'e livlu se.*
 3SG give-3SG book SP
 ‘He gave him the book.’
 b. **Ê da livlu se {ê / êlê}.*
- (21) a. *Ê da Zon êlê.*
 3SG give Zon 3SG
 ‘He gave it Zon.’
 b. **Ê da êlê Zon.*
- (22) a. *Ê d'e êlê.*
 3SG give-3SG 3SG
 ‘He gave it him.’
 b. **Ê d'êlê {ê / êlê}.*

Hence, it can be concluded that DOCs in Santome are in all respects totally static constructions that provide negative evidence for verb movement. Consequently, the simple transitive structure in (18) is able to account for the observed patterns.

3.2.1.4. Allomorphic variation: *ba* and *be* ‘to go’

A more complex piece of evidence against verb movement in Santome comes from the complementary distribution of *be* and *ba*, both meaning ‘to go’, and in particular from the behavior of goal arguments and comitatives.⁵⁶ This pair is unique in the sense that no other allomorphic variation of this type is found in this language. Ferraz (1979: 89) first mentioned this pair and describes *ba* as occurring before locatives and as an auxiliary verb, whereas *be* occurs elsewhere. Some examples of this distribution are found in Table 1.

Table 1. Complementary distribution of *ba* and *be* ‘to go’ (Hagemeijer 2005a: 73).

<i>Ba</i>	<i>Be</i>
<i>Ê ba</i> [ala]. 3SG go there 'He went there.'	<i>Ê be</i> [dai]. 3SG go from-here 'He went from here.'
<i>Ê ba</i> [ke]. 3SG go house 'He went home.'	<i>Ê be</i> [ku bô]. 3SG go with you 'He went with you.'
<i>Ê ba</i> [omali]. 3SG go sea 'He went to the sea.'	<i>Ê be</i> [d'omali]. 3SG go by-sea 'He went by sea.'
<i>Ê ba</i> [wê karu]. 3SG go front car 'He went to the front of the car.'	<i>Ê be</i> [ni wê karu]. 3SG go in front car 'He went in the front seat of the car.'

In Hagemeijer (2000) I argued that this distribution was the result of the positive or negative specification of a telic feature that I claimed stood for an opposition between verbs of the unaccusative and unergative type respectively. But in the light of the data provided in Hagemeijer (2000), Becker & Veenstra (2003) argue that the distinctive forms appear to be the result of morphological encoding of the selection properties of these verbs: *be* occurs with adjuncts; *ba* selects arguments. It will be shown that this claim is essentially correct.

⁵⁶ Cf. Hagemeijer (2005 a: 72-73) for a discussion of the etymology of these allomorphs.

The special status of goals in the argument structure of directed motion verbs follows from the fact that they typically occur adjacent to the verb (*ba*) and from their preposition-less shape. Compare the following pairs which show that goal DPs cannot interchange positions with, for instance, PPs in example (23), aspectual constructions in (24), or participles in (25).

- (23) a. *Ê ba/*be [losa] [ni wê karu].*
 3SG go plantation in front car
 ‘He went to the plantation in the front seat of the car.’
 b. **Ê ba/be [ni wê karu] [losa].*
- (24) a. *Ê ba/*be [Pla Konsa] [ka kôlê].*
 3SG go Beach Shell ASP run
 ‘He went running to the Beach of Shells.’
 b. **Ê ba/be [ka kôlê] [Pla Konsa].*
- (25) a. *Ê ba/*be [ke] [tasondu].*
 3SG go home seated
 ‘He went home seated’
 b. **Ê ba/be [tasondu] [ke].*

The following examples show that implicitly known locations also trigger *ba*. In (26), the location implied is the place where the king is. In (27), the sound of the falling breadfruit expressed by the ideophone *din* identifies the ground as on the endpoint of their fall.

- (26) *...pa non dêsê ba sun alê.*
 ...for 1PL descend go Mr. king
 ‘...in order for us to go down to the king(‘s place).’
- (27) *N kônô dôsu kabêsa ba din.*
 1SG collect two head go IDEOPHONE
 ‘I cut off two heads (of breadfruit) that hit the ground.’

Becker & Veenstra (2003) point out that the telicity analysis is problematic because *be* appears when goal-denoting arguments are interchanged with comitative constituents (cf. 28a-b) and when a Wh-argument is moved to the clause-initial position.

- (28) a. Ê *ba/*be* [ke Zon] [ku inen mina se].
 3SG go home Zon with 3PL child SP
 ‘He went to John’s place with these children.’
- b. Ê *be/*ba* [ku inen mina se] [ke Zon].
 Both: ‘He went with these children to John's place.’
- (29) *Andji ku ê subli be/*ba?*
 where KU 3SG go up go?
 ‘Where did he go up to?’

Informally stated, the data above show that whenever the goal does not occur in a strictly adjacent position to the verb, *be* has to surface.

3.2.1.4.1. Case-marking

Consider the following pair of sentences where a different verb form is triggered despite the fact that the verb occurs with a goal argument:

- (30) *Zon be/*ba antê awa.*
 Zon go until river
 ‘Zon went as far as the river.’
- (31) *Maya ba/*be nglêntu awa.*
 Maya go inside river
 ‘Maya went into the river.’

The difference cannot, of course, be explained by means of telicity, since both sentences have a telic reading. In addition to the different verb selection, the two constructions above also differ with respect to adverb placement. As shown, adverbs cannot occur between a verb and its internal argument. Therefore, an adverb like *ndjandjan* ‘quickly, fast’ is unable to intervene between a directed motion verb and its goal complement in (32), contrasting with cases like (33) and (34).

- (32) Ê *ba* (**ndjandjan*) [*liba ke*] (*ndjandjan*).
 3SG go (quickly) top house (quickly)
 ‘He went on (top of) the house quickly.’

- (33) *Ê be (ndjandjan) [antê poson] (ndjandjan).*
3SG go quickly until city quickly
'He went quickly to the city of S. Tomé.'

The obvious conclusion is that only *ba* has Case-assigning properties, whereas the constituent *antê awa* 'until the water/river' cannot receive Case from the verb. It turns out that this difference can be explained by the way Case-marking relations are encoded in Santome. Despite the traditional view that prepositions are considered items of a closed-class with the categorial label [-N,-V], it is well known that cross-linguistically the 'prepositional function' can be fulfilled by lexical elements from different categories.

In Santome, most items that exhibit a 'prepositional function' cannot be considered prepositions proper. In fact, nouns and verbs fill in this function to a significant extent (cf. Hagemeyer 2005a). The [+N, -V] (nouns) class comprises nominals such as *nglêntu* 'inside'. The items that feature as [-N, +V] (verbs) are typically verbs in the VP₂ slot of serial verb constructions.⁵⁷ The following table illustrates this tripartite categorial system of prepositions proper, nominals and the second verb in serial verb constructions:

⁵⁷ This statement somewhat obscures the complexity of the grammaticalization paths of these items, especially the second verb in serialising constructions (Hagemeyer 2000, 2001). It can be shown that some of these verbs exhibit both prepositional and verbal properties.

Table 2. Lexical items with ‘prepositional functions’.

[-N, -V] → <i>be</i>	[+N, -V] → <i>ba</i>	[-N, +V] (V ₂ in serialization)
<i>di</i> ‘of’	<i>wê</i> ‘in front of, the front, eye’	<i>da</i> ‘for, from, to, because, on’ (to give, to hit) ⁵⁸
<i>ni</i> ‘in, from’	<i>tlaxi</i> ‘behind, the backpart, back’	<i>pê</i> ‘in’ (to put)
<i>antê</i> ‘until’	(<i>n</i>) <i>glêntu</i> ‘inside, the inside’	<i>be/ba</i> ‘to, away’ (to go)
<i>jina</i> ‘from, since’	<i>liba</i> ‘on top of, upper part’	<i>bi</i> ‘from’ (to come)
<i>sê</i> ‘without’	<i>basu</i> ‘beneath, under(neath), lower part’	<i>subli</i> ‘up’ (to go up)
<i>ku</i> ‘with’	<i>ômê</i> ‘centre, middle, between’	<i>loja</i> ‘around’ (to surround, to go around)
	<i>bodo</i> ‘next to, side’	<i>lêlê</i> ‘alongside’ (to follow, to accompany)
		<i>fô</i> ‘from, since’ (to come from)
		<i>vala</i> ‘across’ (to pass)
		<i>kyê</i> ‘in(side), on’ (to fall)
		<i>xê</i> ‘away from, out of’ (to leave)
		<i>lentla</i> ‘inside’ (to enter)
		<i>kaba</i> ‘to finish’
		<i>pasa/vala</i> ‘than’ (to surpass)
		<i>mundja</i> ‘up(right)’ (stop, stand)
		<i>pê</i> ‘in’ (to put)

For the present purpose, I will focus on the items in the first two columns. Crucially, all the items in the first column occur without exception with *be*, whereas all the [+N,-V] items in the second column require *ba*. This contrast is illustrated in the following pair of sentences:

(34) *Ê be {d’omali / antê omali}.*
 3SG go by-sea / until sea.
 ‘He went {by sea / as far as the sea}.’

(35) *Ê ba {wê/nglêntu} ke.*
 He go front/inside house
 ‘He went {in front of / inside} the house.’

⁵⁸ Although I have included *da* in this class, it lacks verbal features in what resemble serial verb constructions and should therefore be integrated into the first column (Hagemeyer 2000).

These examples show that the selection of a non-prepositional goal argument of the second column has an overt reflex in the verb form. Prepositions proper in the first column assign Case to their object in a standard fashion. The nominal items in the second column have to receive Case directly from the verb. Hence it also follows that locative items such as *ala/nala* ‘there’ occurring with *ba* are actually Case-marked nominals.⁵⁹

There is language-internal evidence for the different status of prepositions proper and nominal preposition-like elements. Unlike prepositions proper, all the nominals listed in the second column of Table 2 can be used intransitively, as in (36), whereas the prepositional items in the first column cannot, as in (37).

(36) *Zon ba nglêntu/wê.*

‘Zon went inside/front.’

(37) **Zon be antê/di/jina/ku.*

Zon went until/of/with

Secondly, prepositions such as *antê* are able to select the nominals of the second column of Table 2:

(38) *Zon be* [_{PP} *antê* [_{DP} *liba* [_{DP} *budu*]]].

Zon go until top stone

‘Zon went until on top of the stone.’

Another matter that needs to be settled is how DPs that follow nominal prepositions are case-marked. Consider the DP *ke* ‘house’ in the following example or *budu* in (38) above.

(39) *Maya ba* [_{DP} *nglêntu* [_{DP} *ke*]].

Maya go inside house

‘Maya went inside the house.’

⁵⁹ In some specific cases prepositional *ni* ‘in’ contracts with nominals, for example *nglêntu* ‘inside’ or *nala* ‘there’, and *ba* still occurs. I assume that these items have become reanalysed as a single lexicalized item. If the preposition *ni* had preserved Case-assigning properties, it would be expected to surface with *be*.

Since DPs do not have direct Case-marking properties, the insertion of an additional Case-marking item is required to mediate the relation with another DP, as in the case of English ‘of’. In Santome, the insertion of such an element is not visible at the surface, but becomes clear upon extraction of the relevant argument. This is exemplified by focus and left dislocated constructions, in (40) and (41) respectively, where *di* ‘of’ is obligatorily inserted. The contraction of *di* and spelled-out trace *ê*, signaling that movement has taken place, yields *dê*.⁶⁰

(40) *Awa so Maya ba nglêntu *(dê).*
 river FOC Maya go inside of-3SG
 ‘It was the RIVER Maya went into.’

(41) *Karu, Zon ba wê *(dê).*
 car Zon go front of-3SG
 The car, Zon went to the front of it.

In this section it was shown that the essentially correct argument/adjunct distinction (Becker & Veenstra 2003) can be restated as a more general principle of Case-marking, which has the slight advantage that goal-denoting arguments that do not occur with *ba*, such as the *antê*-construction, can receive a thematic role from the verb but receive Case from the preposition. This principle explains away the bulk of the data on *ba* and *be*.

The data up to this point have further shown that *ba* is empirically restricted to environments where two conditions have to be fulfilled, namely

- i) the presence of an overt or implicit goal-denoting DP;
- ii) the adjacency of this argument to the verb.

Having now set the stage with respect to the properties of this verb pair, the next section will focus on a particular construction, namely the comitative-goal shift and its implications for verb movement.

⁶⁰ Cf. Hagemeyer (2000) and Alexandre & Hagemeyer (2002) for a discussion of spelled-out traces and resumptivization in Santome. Note also that vowel-initial nouns unequivocally show that there is a Case-maker in these structures: *nglêntu *(d)’awa* ‘inside (of) the water/river.’

3.2.1.4.2. The comitative-goal shift

This section addresses why comitatives, unlike other constituents, are able to shift together with the goal argument and how this relates to our understanding of the position of the verb in the clause. Upon detailed investigation of the comitative-goal shift, it will follow that verb movement, even its shortest version, is hardly tenable.

Despite the argumental status of goals argued for above, it was shown that comitatives are exceptional because they can intervene between the verb and the goal, as in (42). With non-goal arguments, this option is precluded, as illustrated in (43).

(42) \hat{E} *be* [ku migu dê] [ke Zon].
3SG go with friend POS house Zon
'He went with his friend to Zon's place.'

(43) * \hat{E} *kume* [ku migu dê] [pixi].
3SG eat with friend POS fish
'He ate fish with his friend.'

Note, in the first place, that comitatives are always optional. Yet, concomitant constituents play a special role in argument structure because they are linked to another participant (the subject in the cases under discussion), although this does not necessarily entail equal participation. Cross-linguistically, the concomitant relation manifests itself in several domains, from Theta-sharing to (less common) instances of Case-sharing (cf. Lehmann & Shin 2005).

In spite of their specific status, comitatives are generally not considered to be arguments. Baker (1992), for instance, considers comitatives to be non-subcategorized second agents or second themes, i.e. constituents lacking a *primitive* thematic role. This does not necessarily imply that comitatives also behave like adjuncts. Schütze (1995), for instance, concludes that in English instrumentals, and comitatives alike, behave more like arguments.

In Santome, goals and comitatives share several properties, among which I would like to highlight the acceptable extraction from NP-islands of a D-linked Wh-constituent, as shown in (44a-b).

- (44) a. *Kê mosu ku Zon kunda ku mwala ku ska dwêntxi be*
 What boy KU Zon think that woman REL ASP ill go
ku ê?
 with 3SG
 ‘What boy did Zon think that the woman who is ill went with?’
- b. *Kê fela ku Zon kunda ku mwala ku ska dwentxi be?*
 What market KU Zon think that woman REL ASP ill go
 ‘What market did Zon think that the woman who is ill went to?’

The basic difference between these two constituents relates to adverb placement. Comitatives are more flexible with respect to the position in which they can occur. It was shown that typical VP-adverbs could not intervene between the verb and the goal (cf. 44a). This, however, is fully acceptable with comitative PPs (cf. 44b).

- (45) a. *Zon ba (*ndjandjan) fela (ndjandjan).*
 Zon go (quickly) market (quickly]
 ‘Zon went to the market quickly.’
- b. *Zon be (ndjandjan) ku anzu (ndjandjan).*
 Zon go (quickly) with baby (quickly)
 ‘Zon went quickly with the baby.’

Note further that, unlike comitatives, typical VP-adverbs that are also introduced by preposition *ku* ‘with’ cannot be stacked between the verb and the goal, as illustrated in (46) and (47).

- (46) *Zon ka lentla (*ku ope dê) palaxu *(ku ope dê).*
 Zon ASP enter (with foot POS) palace (with foot POS)
 ‘Zon enters the palace by himself.’
- (47) *Zon ba (*ku fomi) xipitali *(ku fomi).*
 Zon go (with hunger) hospital (with hunger)
 ‘Zon went hungry to the hospital.’

The examples in (48) below further show that goals cannot be separated from the verb by more than one constituent (cf. 48a). Example (48b) shows that there is no rigid ordering between VP-adverbs and the comitative when they follow the goal.

- (48) a. *Zon be ku mwala (??ndjandjan) fela (ndjandjan).*
 Zon go with woman (quickly) market (quickly)
 ‘Zon went with the woman to the market quickly.’
- b. *Zon ba fela (ndjandjan) ku mwala (ndjandjan).*
 ‘Zon went (quickly) to the market with the woman (quickly).’

This difference confirms the intuition that despite their special status comitatives are best analyzed as adjuncts, and goals as arguments.

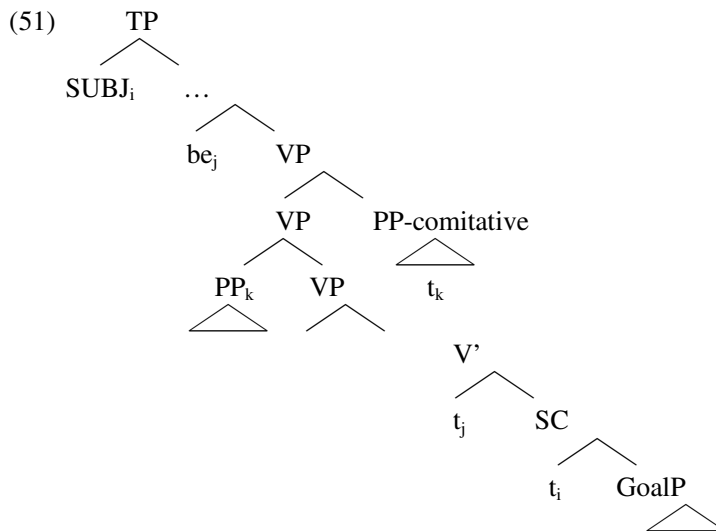
The reason behind the shift is related to the information structure of the sentence. New information, i.e. the questioned material, occurs in the right periphery of the sentence, as shown in the following pairs. Therefore, the answers in (49b) and (50b) are appropriate with respect to the questions in (49a) and (50a) respectively, whereas (49c) and (50c) are not.

- (49) a. *Kê ngê ku Zon ba ke ku ê?*
 what person KU Zon go house with 3SG
 ‘With whom did Zon go home?’
- b. *Ê ba ke ku inen mina se.*
 3SG go house with PL children SP
 ‘He went home with these children.’
- c. *??Ê be ku inen mina se ke.*
- (50) a. *Andji ku Zon be ku inen mina se?*
 where KU Zon go with PL child SP
 ‘Where did Zon go with these children.’
- b. *Ê be ku inen ke.*
 3SG go with 3PL house
 ‘He went with them home.’
- c. *??Ê ba ke ku inen.*

This means that the comitative PP typically has the status of old information when it precedes the goal. All the instances of pronominalized comitatives (either animate or non-animate) in my corpus occur to the immediate right of the motion verb, which is predicted from the fact that pronouns typically have old informational status. Non-pronominalized comitatives can of course occur to the left or the right of the goal according to their informational status.

3.2.1.4.2.1. Scrambling and verb movement

Structurally speaking, one could propose a scenario in which the comitative is basically right-adjoined to VP and left-adjoins through scrambling to VP after the verb has moved out of VP. Scrambling would be motivated by the need to escape from the default clause-final focus position, following Reinhart (1995). This scenario is illustrated in the following tree structure:



Note that I assume that the presence of a goal corresponds to an unaccusative structure. The subject and the goal form a small clause (SC) as proposed by Hoekstra & Mulder (1990). I accordingly treat *ba/be* as unaccusative verbs. Although one of the uses of *be* is intransitive, it is typically perfective in the sense that it focuses on the movement away of the deictic point of reference. This becomes clear by adding an event-delimiting adverb, as in example (52), which distinguishes manner of motion verbs from directed motion verbs.

- (52) *Zon be/bi/*kôlê n'ũa minutu.*
 Zon go/come/run in-one minute
 'Zon went/came/*ran in a minute.'

Note also that the comitative is right-adjoined to VP in the spirit of Ernst (2002). After construing the VP, the subject moves in a standard fashion to [Spec,TP] and the verb would arguably raise and adjoin to an aspectual node, given the strict adjacency between Asp° and V°. After verb movement, the comitative PP would be merged as a left-adjunct to VP to derive the correct surface order for S-V-PP_{comitative}-Goal.

There are several problems with this hypothesis, though. First, comitatives cannot scramble with arguments that are not goals, as follows from a comparison of (42) and (43). Second, it is not clear why only comitatives - and not other adverbials - would be able to scramble and left-adjoin to VP. Moreover, basic left-adjunction to VP is not allowed at all. Here I follow Costa (1998: 288), who suggests that adjunction by movement cannot target a category where base-generated adjunction is impossible. Third, the motivation for verb movement under this hypothesis is rather obscure and seems to hinge exclusively on deriving the correct order with comitatives.

In addition to the comitative, it should be noted that there is another case that breaks up the surface adjacency of the verb and the goal, namely pseudo-reflexive pronouns (section 2.6.3). Whenever the clause contains a pseudo-reflexive (PSR), a comitative and a goal, the goal obligatorily precedes the comitative (cf. contrast between examples (53a) and (53b)).

- (53) a. *N be mu poson ku piskadô.*
 1SG go PSR city with fisherman
 'I went to the city of São Tomé with the fisherman.'
- b. **N be mu ku piskadô poson.*
 (I went with the fisherman to the city of São Tomé)

Note further that it was shown in section 2.6.3 that these forms are underlying PPs and always trigger *be*. An analysis of scrambling and verb movement is thus unable to satisfactorily account for the data and is counterintuitive with respect to the linguistic properties of Santome.

3.2.1.4.2.2. DP-adjunction

Having discarded both scrambling of comitatives and verb movement for language-internal reasons, there are at least two analyses that make better predictions with respect to the data: i) adjunction to the DP with which the comitative is in a concomitant relation or ii) right-adjunction to VP (discussed in section 3.2.1.4.2.3 below).

The first hypothesis follows the analysis of Ionin & Matushansky (2002) who provide a unified analysis of Russian comitatives, which, in their view, may hold for other languages as well. Under this proposal, the different positions in which comitatives are found are either a reflex of extraposition or stranding. Despite the interest of this analysis, which derives especially from the fact that concomitance is an underlying local relation between participants, it fails to explain the following facts in Santome.

If the comitative is adjoined to the subject of unergative/transitive clauses which are generated in [Spec,VP] in the usual fashion, the comitative precedes the verb on the surface. Therefore either the verb has to move, with all the consequent problems, or the comitative has to be obligatorily extraposed. However, extraposition runs into the problem that the comitative can occur between the goal and a typical VP adverb such as *ndjandjan* ‘quickly’. This is unexpected because after building the VP, extraposition would target a VP-final slot.

Furthermore, comitative preposition *ku* is also used for DP-coordination, as in *Zon ku Maya* ‘Zon and Maya’, which, despite exhibiting the same preposition, should arguably receive a different syntactic structure than the comitative. Since the DP-adjunction hypothesis is a strong hypothesis in the sense that it is meant to furnish a single structure for all comitative *ku*-phrases, coordination above would be the result of moving the DP+PP from a VP-internal position to the surface subject position. However, true concomitance would be the result of splitting the VP-internal DP+PP and moving the DP and PP to its respective surface positions.

Finally, pseudo-reflexives also constitute counterevidence to the DP-adjunction hypothesis. This basically follows from the discussion in section 2.6.3, where it was argued that pseudo-reflexives are phonological clitics that force the Goal to extrapose. So, under the DP-adjunction hypothesis, the comitative would have to be extraposed for the same reasons as the goal but this is a counterintuitive solution and faces the problem that comitative-stacking between the pseudo-reflexive pronoun and the goal is not

possible. Hence, the arguments above make comitative adjunction to DP an undesirable solution.

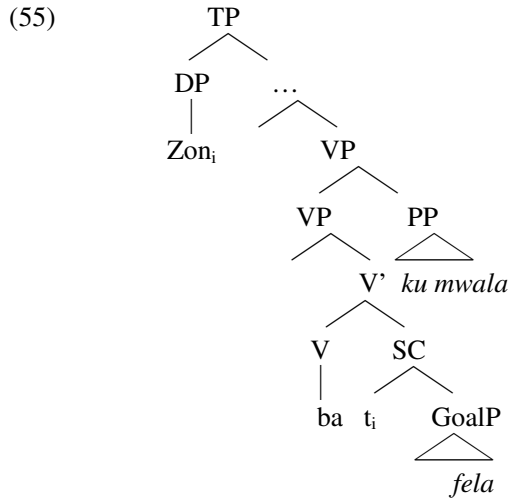
3.2.1.4.2.3. Right-adjunction to VP

Right-adjunction of the comitative to VP, on the other hand, is fairly unproblematic and makes good predictions with respect to the different word orders in transitive/intransitive clauses. When the comitative precedes the goal, I assume that the goal is extraposed, right-adjointing to VP, where it receives focus. The comitative then becomes automatically defocused.

I am, however, aware of a single problem that also arises under the DP-adjunction hypothesis, namely the impossibility of extraposing the goal (i.e. the comitative in the other hypothesis) to the final position when there are two VP-adjuncts:

- (54) a. *Zon be ku mwala fela ndjandjan.*
Zon go with womanmarket quickly
'Zon went with the woman to the marker quickly.'
- b. ??*Zon be ku mwala ndjandjan fela.*
- c. ?? *Zon be ndjandjan ku mwala fela.*

Assuming that the comitative and the VP adverb are right-adjoined, extraposition is expected to follow the adverb, which results in a degraded sentence for most speakers I have consulted. However, note that in (54a) the adverb in final position has to be prosodically marked, which suggests that it is only adjoined after extraposition of the goal or, alternatively, that there is post-syntactic reordering going on at PF. This is particularly appealing in a language with a rigid syntax above V (i.e., no verb movement, base-generated TMA-markers, etc.). Therefore, the final solution that I propose for a sentence where the goal precedes the comitative now is shown in the following representation:



This structure represents the derivation of an unaccusative predicate. If the GoalP *fela* ‘market’ is extraposed, I assume that it right-adjoins to the comitative VP. Thus, the discussion of the allomorphs *ba* and *be* proved to be an important tool for dismissing the need for any sort of verb movement in Santome. Moreover, it was also shown that it is not desirable to preclude right-adjunction from Santome’s syntactic repertoire.

3.2.1.4.3. *Ba* and *be* and the lexicon

To finalize the discussion on *ba* and *be*, I will briefly address how the lexicon and syntax deal with these forms. Here it is crucial to remember that there are many structures (cf. Table 1) that do not exhibit a goal argument. Invariably, *be* occurs in these cases. An important question is then, of course, whether *be* always has a goal in its argument structure. The following examples illustrate that this is arguably not the case:

- (56) *Zon be dai.*
 Zon go from-here
 ‘Zon went around.’ / ‘Zon went from here.’
- (57) *Zon be dê.*
 Zon go PSR
 ‘Zon went away.’
- (58) *Zon kôlê be/bi.*
 Zon run go/come
 ‘Zon ran away/Zon came running

In these examples, the focus is on the movement away from the deictic centre and there is no implication of a goal argument. Hence I assume that ‘to go’, and arguably the limited range of other verbs of directed motion as well, can be treated as a transitive or an intransitive verb according to the construction in which they occur. I consider transitivity to be an unspecified feature in the lexicon.

A clearer picture now starts to emerge. It followed from the distribution of *be* that this allomorph occurs in intransitive and transitive environments. As for *ba*, it only occurs in transitive constructions and under the condition that there is a goal DP adjacent to the verb. Table 3 sums up these findings.

Table 3. Distribution of *be* and *ba*.

- Transitive	+ Transitive	
	<i>Non-adjacency to goal DP</i>	<i>Adjacency to goal DP</i>
<i>be</i>	<i>be</i>	<i>ba</i>

The conclusion is, therefore, that *be* should be considered the default form. Historically this also makes some sense, since it was shown that *be* better complies with the phonological rules applied to the Portuguese lexicon. Moreover, all GGCs exhibit *be*, but Lung’iye lacks *ba*. Becker & Veenstra (2003) argue that the change from *ba* to *be* is determined post-syntactically, a view that finds support in the different forms that are related to movement operations, such as Wh-movement (cf. 59), but also to Focus constructions (60) or goal extraposition, in (61).

(59) *Andji ku Zon be?*

where KU Zon go
‘Where did Zon go?’

(60) *Losa so ê be.*

plantation FOC 3SG go
‘It was to the plantation he went.’

(61) *Zon be ku migu fela.*

Zon go with friend market
‘Zon went with a friend to the market.’

My claim is that, although *be* and *ba* are in the lexicon, *be* is the default form drawn from the lexicon and standardly merged into the structure of intransitive and transitive directed motion predicates. At spell-out, *be* is pronounced, unless the requirements for *ba* are met (goal argument+adjacency). In this sense, *ba* is then the post-syntactic verb form.

3.2.2. Summary

In the previous sections I argued that Santome does not exhibit verb movement and that it therefore matches with what has been claimed for most creole languages. In addition to two standard tests, namely adverb placement and floating quantifiers, the existence of DOCs and the distribution of comitatives and goals with respect to the allomorphs *be* and *ba* show that verbs remain *in situ*. The absence of inflectional morphology and the presence of lexicalized functional projections in the extended VP, to be discussed in the remainder of this chapter, yield a strongly isolating language that meets the expectation that verbs do not move and adjoin to higher projections.

3.3. The TMA-system

Since Bickerton (1981), TMA-systems have been at the core of creole studies because of the alleged similarities between these systems across the whole range of creole languages and the universal implications of this claim. Even though Bickerton's strong version of the Language Bioprogram Hypothesis could not be upheld in the light of new findings, often inspired by the program itself (cf. Veenstra *forthc.*), for a long time TMA-marking was an island of resistance to superstrate and substrate views on creolization. Outside creole studies, cross-linguistic studies on tense, mood and aspect are also abundant and are often conducted in very different theoretical frameworks, as pointed out by Sasse (2002).

Given the cross-linguistic importance of TMA-marking and its special place within creole studies, it is therefore not surprising that this topic has deserved particular attention in the GGC, albeit with a strong focus on tense and aspect. In fact, the earliest attempt to account for tense and aspect in Santome goes back as far as Negreiros (1895), where it is actually the only domain of grammar granted a systematic account. Other studies of tense and aspect, and sometimes a glance at mood, in the GGC can be found

in Barrena (1957), Valkhoff (1966), Günther (1973), Ferraz & Valkhoff (1975), Ferraz (1979), Muysken (1981), Bickerton (1981), Holm (1989), Post (1995), Maurer (1995, 1997), Lorenzino (1998) and Schang (2000). Only some of these studies rely on independent fieldwork.

Bickerton (1981:76) himself critically reviews the, at that time, earlier work on TMA-markers in Santome, namely Valkhoff (1966), Ferraz (1979) and Muysken (1981), stating rather skeptically that “[...] the reader had better not even attempt to follow the names which the various tenses, modes, and aspects are given by these three authors.” After Bickerton (1981), there is a long period of silence with respect to TMA-marking in Santome. This topic is reopened by Schang (2000: 181-201), who provides a synthetic analysis that is essentially based upon the basic tense and aspect patterns. Apart from these studies on Santome, there are more in-depth discussions of TMA-marking in Ngola (Maurer 1995), Lung’ie (Maurer 1997, *forthc.*) and Fa d’Ambô (Post 1995). Not so surprising, the empirical findings of these studies are to a significant extent applicable to Santome, as explicitly mentioned by Maurer (1995, 1997).

This sub-chapter is structured as follows. First, I will briefly summarize earlier proposals. Next, I will survey the core tense and aspect markers and their combinations. This is followed by section 3.3.3 on stativity and section 3.3.4 on relative tense. Sections 3.3.5 and 3.3.6 are concerned with the syntactic representation of aspect and tense respectively. In section 3.3.7 the findings on tense and aspect are summarized. Finally, section 3.3.8 proposes that Santome exhibits two functional projections to encode mood and modality.

3.3.1. Earlier proposals on tense and aspect in Santome

The topic of tense and aspect has been popular in studies on the GGC. Table 4 below sums up the classification of the tense-aspect markers proposed by Negreiros, Valkhoff, Ferraz and Holm. I have unified the orthography of the forms and translated Negreiros’ labels from Portuguese to English. The final column shows whether the marker is attested in contemporary Santome.

Table 4. Classification of tense and aspect markers in the GGC.

	<i>Negreiros</i> (1895:324-34)	<i>Valkhoff</i> (1966: 104-13)	<i>Ferraz</i> (1979: 81-4)	<i>Holm</i> (1989: 149)	<i>Attested in contemporary sources</i>
$\emptyset + V$	-	preterite	perfect	unmarked	✓
$\emptyset + V + za$	perfect	perfective	-	completive anterior	✓
<i>ka + V</i>	present	habitual, future	habitual, future	habitual	✓
<i>tava ka~ta ka + V</i>	imperfect	past habitual	past progressive	habitual/progressive anterior	✓
<i>sa ka~ska~xka + V</i>	-	progressive	present progressive	progressive	✓
<i>tava + V</i>	pluperfect	past completive	pluperfect	unmarked anterior	✓
<i>ka bi + V</i>	future imperfect	incompletive future	-	-	✓
<i>kia + V</i>	imperfect	expectative	(prospective) ⁶¹	-	✓
<i>tê + Past Participle</i>	composite perfect	-	-	-	-
<i>tê d(j)i bi + V</i>	composite future imperfect	-	-	-	✓
<i>tê d(j)i + V</i>	composite future perfect	completive future	-	-	✓

It follows that, despite some terminological differences, there is a reasonable overall degree of consensus regarding the core tense-aspect system of Santome.

Negreiros, the pioneer in this domain, clearly follows the Portuguese classification of tenses used by grammarians of his time. The construction *tê*+Past Participle is not attested anywhere else and some of his labels seem to lack descriptive adequacy, for instance with respect to *kia*, which might ultimately be due to the tradition in which his work is couched. Note as well that Negreiros did not include the progressive in his description, although it can be found on a few occasions in his examples.

Valkhoff (1966) considers Santome a language based on aspect and not on tense and establishes a primary distinction between completive and incompletive (i.e. perfective and imperfective). In addition, he mentions that verbs like *sêbê* ‘to know’, *tê*

⁶¹ Ferraz attests the form *kia*, which he translates as ‘to be about to do, to nearly do’, but doesn’t provide a label for it.

‘have’ and *sa* ‘to be’ do not take what he labels the prefix *ka* because “the action is expressed or completed once and for all (1966: 106)”.⁶² What is of course meant by his words is that these verbs are statives and behave differently from non-stative verbs, as will be shown in detail in section 3.3.3.

Ferraz (1979) essentially follows Valkhoff’s proposal, with a few minor differences. Valkhoff only uses the contracted form *ska*, whereas Ferraz lists the contracted form and the non-contracted form *sa ka*. Moreover, Ferraz’s past progressive does include the past progressive but it also includes the past habitual, although he does not explicitly use this label.

I did not include Schang (2000) in the table above because the goal of his discussion of the tense-aspect markers in Santome is to establish the core constraints imposed on the markers, as per Table 5.

Table 5. Value of TMA-markers according to Schang (2000: 197).

<i>Markers</i>	<i>Interpretative constraints</i>
∅	contextual + conceptual constraints (encoded by the verb)
<i>ka</i>	imperfective
<i>sa</i>	present
<i>tava</i>	anterior

Schang proposes the following structure of Santome tense-aspect:

(62) [tense _[present, anterior] [aspect _[imperfective] [verb]]]

In his view, it cannot be claimed that the unmarked (or ∅-marked) forms of the verb are perfective because there are several environments that do not support this hypothesis (e.g. stative verbs and conditional clauses). According to this author, the label imperfective for *ka*, for instance, covers present tense, future tense, as well as counterfactuality. Schang also departs from previous analyses (Negreiros, Valkhoff and Holm) by explicitly and correctly arguing against clause-final adverb *za* ‘soon, already’ as a part of the core tense-aspect system.

⁶² Schuchardt (1888: 25), who makes very few observations about TMA in Santome, was actually the first to note that bare forms may also refer to the present.

All the above authors implicitly or explicitly mention that one of the functions of *ka* is to mark irrealis, for instance conditionals, but fail to identify the highly specific phono-syntactic properties of this marker, which accordingly should receive a different treatment (cf. section 3.3.8.1).

Finally, the behaviour of tense and aspect markers with different verb classes, i.e. stative and dynamic verbs, is hardly explored in a systematic way in the literature on Santome. Maurer (1995, 1997), however, has addressed this issue extensively in the case of Lung'ie and Ngola and suggests that his findings apply to Santome as well. I will address his proposal in section 3.3.3. In fact, the detailed descriptions of tense and aspect in these two related creoles are fundamental to our understanding of their workings in Santome.

In section 3.3.2, I will provide an overview of core tense and aspect markers and how they interact. From a descriptive point of view, it will be shown that the system is actually much richer and much more complex than follows from the existing literature on this topic.

3.3.2. Tense and aspect markers

As has been pointed out abundantly in the literature mentioned in the previous section, the core of the grammatical tense and aspect system of Santome consists of the following preverbal items: (i) *ka~ga*⁶³, (ii) *sa ka~ska~xka*, (iii) *tava~ta* and (iv) the zero marker or non-marked verb. In this section I will provide a descriptive overview of the basic functions of tense and aspect markers and the way they combine. I will be mostly focusing on the combination between these markers and the class of dynamic verbs. Sections 3.3.3 and 3.3.4 will then provide a more in-depth description of the particularities of the restricted class of stative verbs and show that Santome is a relative tense language. In sections 3.3.5 and 3.3.6, I will discuss the syntax of aspect and tense taking into account the descriptive findings in the next sections.

⁶³ I refer the reader to section 2.6.1 for *ga*, the phonologically conditioned variant of *ka*.

3.3.2.1. *Ka*

This marker is by and large the most common in the Santome TMA system. Ferraz (1979) claims that *ka*, and its phonologically conditioned variant *ga*, has three functions: it marks habitual aspect, in (63), future tense⁶⁴, in (64), and conditionality, in (65).

- (63) ...*so* *bô ka bêbê ãa bon vinpema doxi.*
...then 2SG ASP drink one good palm wine sweet
'...then you drink a good, sweet palm wine.'
- (64) *Sabino pô môlê; a ka lembla Sabino.*
Sabino may die; IMP ASP remember Sabino
'I may die, but I will be remembered.'
- (65) *Ami za n gá sa mama, n gá kunga klusupê xi-xinku.*
1SG already 1SG SBJV be lady 1SG SBJV put cross put RED-five
'If I were the lady, I would have put the cross in all the five boxes.' (on the ballot paper)

As illustrated, the first two functions are in fact well attested, but conditional *ká* (*gá*) differs from *ka* not only by bearing a high tone, but also because it occurs in a different syntactic position, which can be seen in counterfactual conditionals (cf. section 3.3.8.1). Furthermore, the use of *ka* as a future marker is typically determined by context or by the presence of temporal operators, such as *amanha* 'tomorrow' and *ola* 'when' in the following examples.

- (66) *Amanha n ga lema ngêmbu.*
Tomorrow 1SG ASP trap bat
'Tomorrow I will trap bats.'
- (67) *Ola n ga kôlê kabêsa, n ga lembla Didi ê.*
when 1SG ASP run head 1SG ASP remember Didi EMPH
'When it runs through my mind, I remember Didi.'

⁶⁴ More specifically, *ka* is a definite future marker in the sense of Lefebvre (1998), meaning that the event in question will indeed take place somewhere in the future.

Depending on the type of predicate, *ka* may be interpreted as an inchoative marker:

- (68) *Mosu ka dwêntxi.*
boy ASP fall ill
'The boy is falling ill.'

This marker can further be interpreted as a progressive in specific environments, such as the perception construction (69), the past imperfective construction (70), certain auxiliary constructions with aspectual and modal verbs (71-72), in imperatives (73) and in enumerations that provide a durative reading of the event (74).

- (69) *Namplakata, pya Mosu Kaximbu ala ka bi.*
suddenly look boy pipe there ASP come
'All of a sudden, look there's the Boy with the Pipe coming our way.'

- (70) *Sangê tava ka dumini za ô!*
she was ASP sleep already EMPH
'She was already sleeping!'

- (71) *Ê bila ka fe axi en.*
3SG turn ASP do this way EMPH
'He did this again.'

- (72) *Punda dedu pô ka fede, bô na ka kot'e buta fa.*
because finger may ASP smell2SG NEG ASP cut-3SG throw NEG
'Just because a finger smells, you don't cut it off.'

- (73) *Ka têndê ûa kwa ê!*
ASP listen one thing EMPH
'Listen here!'

- (74) *Tatalugwa sa ka nda, ka nda, ka nda.*
Turtle ASP walk ASP walk ASP walk
'Turtle is walking and walking and walking.'

In all but the past imperfective construction, progressive *ska* can also be used, which shows that there the functions of *ka* and *ska* often overlap. In section 3.3.5 I will argue that there is some evidence that *ska* might be specializing in the progressive functions of *ka*, for instance the *sa ska*-construction (section 3.3.2.8).

3.3.2.2. *Sa ka ~ ska ~ xka*

Ferraz considers these forms free variants and classifies them as the present progressive (1979: 82), which is in fact the main function of this item (75). However, the progressive marker also functions as an aspectual linker between predicates (76) and, like *ka*, may occur in certain auxiliary constructions with aspectual and modal verbs (77-78).

- (75) *Bô ska kia ku ngê tamen.*
2SG ASP raise with people adult
'You are being raised by adults.'
- (76) *Ê mundja sa ka dispidji di pôvô.*
3SG stop ASP say goodbye from people
'He stopped to say goodbye to the people.'
- (77) *Maji n na bila ska lembla fa.*
but 1SG NEG turn ASP remember NEG
'But I didn't remember it anymore.'
- (78) *Kê tipu minjan ku sun Americo ka pô ska bolo?*
what type remedy KU mister Américo ASP can ASP rub
'What kind of remedy is Mr. Americo able to rub?'

Although Ferraz captured the most common function of *ska*, he failed to address at least two situations in which the label "present progressive" doesn't seem to apply. First, there are numerous cases where *ska* receives a past interpretation if past tense is appropriately anchored in previous discourse. In the following example, the bare verbs with a perfective interpretation in the first sentence frame *ska* in the past.

- (79) *Ê d'e kwa kume, ê pê lôpa d'e, p'ê bixi. Mosu*
3SG give-3SG thing eat 3SG put clothing give-3SG for-3SG dress / boy
ska dispidji di mana.
ASP say goodbye from sister
'He gave him food, he provided clothing of his for him to wear. The boy was saying [lit. is saying] goodbye to his sister.'

Second, the progressive marker often has the interpretation of a present perfect.

(80) *Mosu {sa ka / ska} kume pixi.*

Boy ASP eat fish

‘The boy is eating fish.’

‘The boy has been eating fish.’

Although traditionally considered aspect, the present perfect is on the borderline between aspect and tense (cf. Comrie 1976). Each of these readings can be highlighted in the presence of temporal adverbial material, such as *djina onten* and *mê dja* in the following examples:

(81) *Djina onten, mwala {sa ka / ska} dansa.*

since yesterday woman ASP dance

‘Since yesterday, the woman has been/is dancing.’

(82) *Mê dja, mwala {sa ka / ska} dansa.*

noon woman ASP dance

‘At noon the woman is dancing.’

It is a well-known property of progressives that they have the ability to transform dynamic verbs into states (e.g. Moens 1987). The atelic nature of the progressive is confirmed by a number of classical stativity tests identified in Dowty (1979).⁶⁵ For instance, the progressive marker, unlike *ka*, cannot occur in imperatives (83), it contains perfective clauses (84), and it cannot occur in progressive sentences itself (85).

(83) **Ska tëndê ãa kwa ê.*

ASP listen one thing EMPH

(84) *Non ska fisa poto, ola non tëndê glita.*

1PL ASP close door when 1PL hear scream

‘We were closing the door, when we heard screaming.’

(85) **Mwala {sa ka sa ka / ska ska} dansa.*

woman ASP dance

⁶⁵ For a discussion of these tests in relation to Portuguese, I refer the reader to Cunha (1998).

In sum, *sa ka~ska* functions as the imperfective marker for progressive aspect.

3.3.2.3. Non-marked verb

As in many creole languages and beyond, the particularly common zero marker in simplex sentences typically marks the past perfect of non-stative verbs, as in (86), and the present tense of statives, exemplified in (87).

- (86) *Ol'ê bila, sun fisa poto.*
when-3SG turn, man close door
'When he turned around, the man closed the door.'
- (87) *Nansê sêbê dja ku n ga môlê ô?*
2PL know day REL 1SG ASP die EMPH
'Do you know the day I'll die?'

Note, however, that the non-marked verb is not restricted to past perfect readings, because it occurs, for instance, in conditional clauses, complement clauses, finite clauses, in (88-89) respectively.

- (88) *Xi bô na kopl'e fa, bô ka fika sê êlê.*
if 2SG NEG buy-3SG NEG 2SG ASP remain without 3SG
'If you don't buy it, you won't have it.'
- (89) *Ami so ska fada bô pa bô tason nai.*
1SG FOC ASP tell 2SG for 2SG sit down here
'It is me who's telling you to sit down here.'
- (90) *Kont'e da non pa non têndê pikina fan!*
tell-3SG give 1PL for 1PL hear little bit EMPH
'Tell it to us so we can hear a little bit of it!'

Note that these non-marked verbs typically occur in dependent clauses, which are often subjunctive domains, and thus call for a separate treatment which is beyond the scope of this work.

3.3.2.4. *Tava ~ ta*

In agreement with Ferraz (1979: 83), I assume that *tava* is a pluperfect marker for dynamic verbs, as in (90), and a past imperfective marker for stative verbs, as in (91), corresponding to the anterior (past-before-past) marker in the terminology of Bickerton (1981), for whom this cross-creole feature was one of the outstanding creole features of the bioprogram.

- (91) *Bô naxi tava nansê ten fa.*
2SG not-yet TNS born also NEG
'You had not been born yet either.'
- (92) *N ta mêsê pa bô fl'e.*
1SG TNS want for 2SG say-3SG
'I wanted you to say it.'

Tava and *ta* are in free variation. In section 3.3.6 I will argue that *tava* is a marker with the primary feature [Past].

3.3.2.5. *Tava ka ~ ta ka*

Classified by Ferraz (1979: 82) as representing past progressive tense-aspect, this complex past imperfective marker complex for the class of dynamic verbs is used for both past habitual and past progressive.

- (93) *Ûa ja, ami ku kompa mu sun Me Jingu, non ta ka dê sê*
One day 1SG with mate POS Mr. Me Jingu 1PL TNS ASP descend
ba poson.
go town
'One day, me and my mate Mr. Me Jingu, we were going down to town.'
- (94) *Ê tê ãa manu dê ku ta ka bi ai Lomba me.*
3SG have a brother POS REL TNS ASP come here Lomba right
'He has a brother who used to come right here to Lomba's place.'

The form **tava ska* is not attested in contemporary Santome nor in the available sources.

3.3.2.6. *Ka sa ka~ka ska*

This form, which combines two aspectual markers, has not been described in the relevant literature, but is also found in Ngola (*ka thêka*) and Lung'ie (*ka sa*), as noted in Maurer (1995, 1997). According to this author, it has a progressive future reading in Lung'ie and a future imperfective, a habitual imperfective or an irrealis reading in Ngola. As it turns out, the findings for Santome are similar to those of Ngola.

It can be observed from the examples below that the semantics of the individual markers are still visible in the complex form. Thus *ka ska* and its rare non-contracted form *ka sa ka* may express future progressive, in (95), and a habitual progressive, in (96), often with an iterated reading associated with it. Note that this construction also can be used for imperatives, as per (97).

- (95) *Ola ku ê ka sa ka dumini, n ga manda tom'e pê n'ûa tlen.*
When KU3SG ASP ASP sleep 1SG ASP order take-3SG put in-a cart
'When he is asleep, I'll have him put in a cart.'
- (96) *Kada vê ku ê ka ska kanta, mulu ka ka subli ba liba.*
Every time KU 3SG ASP ASP sing wall ASP ASP rise go up
'Every time he's singing, the wall becomes higher.'
- (97) *Ka ska glita, ka glita.*
ASP ASP scream ASP scream
'Keep on screaming and screaming.'

As follows from the examples above, *ka ska* is most common in embedded clauses and in particular temporal and proportional clauses. Nevertheless, my corpus also provides instances of this construction in matrix clauses:

- (98) *Ê ka ska da vin se novu-novu.*
3SG ASP ASP give wine SP young-young
'It [the palm tree] will be giving very fresh palm wine.'
- (99) *Ê kônsê tudu oso. Ê ka ska ndika mu: (...)*
3SG know all bone 3SG ASP ASP indicate me
'He knew all the bones. He used to be always indicating me: (...)'

Just like the progressive marker, this complex form is compatible with a real present reading, in (100), but fails the test for progression that started at a given point in the past, as illustrated in (101).

(100) *Mê dja Zon ka ska dansa.*

noon Zon ASP ASP dance

‘At noon, Zon is/starts dancing.’

(101) **Djina onten, Zon ka ska dansa.*

since yesterday Zon ASP ASP dance

Since it was shown that *ska* is fully compatible with this temporal adverbial, the prediction is that the ungrammaticality of (101) is related to the aspect marker *ka*. Testing *ka* in the presence of an adjunct representing a starting point in the past shows that there is, in fact, an incompatibility.

(102) **Jina onten, ê ka dansa.*

Since yesterday 3SG ASP dance

Thus, *ka* is restricted to present and future contexts, as shown in section 3.3.2.1. For *ka* to refer to a past habitual event, it has to occur in the presence of *tava* (*tava ka*).

3.3.2.7. *Ka ka*

Much like *ska ka*, the aspectual complex *ka ka* is a progressive habitual that indicates the (iterative) aspect of an ongoing stretch of time (103-104), which follows especially from its preference for temporal constructions. Imperatives also occur with *ka ka*, as in (105). In these cases, it can thus be assumed that *ka ka* and *ka ska* share the same semantics.

(103) *Kada vê ku ê ka ka loja ku inen ni tudu yeta se, tudu*

Each time KU 3SG ASP ASP hang out with 3PL in all place SP all

bwê se bila-bila-bila pekadô.

cow SP turn-turn-turn man

‘Each time he’s going about in all those places, all the cows turn into human beings.’

- (104) *Kada vê mina ka ka bila mlagu so.*
 each time girl ASP ASP turn thin only.
 ‘The girl is becoming thinner and thinner all the time.’
- (105) *Ka ka glita, ka glita. Xi ê tënd’ê, bô sêbê za.*
 ASP ASP scream ASP scream if 3SG hear-3SG 2SG know already
 ‘Keep on screaming and screaming. If he hears it, you know what will happen.’

Some speakers share the intuition that *ka ka* is more effective for the imperative function than *ka ska*. On the other hand, *ka ka* additionally insists on the action that is being performed.⁶⁶ Note the contrast between the sentences in (106-108) and in (103-105) above:

- (106) *Sun na tê oji di ka ka piji plaga da ngê fa.*
 you NEG have right to ASP ASP ask curse give people NEG
 ‘You don’t have the right to insist on asking a curse to come down on the people.’
- (107) *Mosu ka ka dwêntxi.*
 Boy ASP ASP fall ill
 ‘The boy is repeatedly falling ill.’ (i.e. a fragile person)
- (108) *A sa nê ka ka futa.*
 IMP be in-3SG ASP ASP steal
 ‘They are stealing and stealing.’

⁶⁶ Günther (1973: 74) notes an identical reduplicated form in Lung’ie (see comment by Maurer 1997: 417), which also exhibits a repeated progressive marker, *sa sa*.

- (i) *Ubudu se ka ka kyê.* (Lung’ie)
 stone SP ASP ASP fall
 ‘This stone may fall at any moment.’
- (ii) *Ubudu se sa sa kyê.* (Idem)
 Stone this be be fall
 ‘This stone is falling right now.’

It follows that, rather than (i), which reads as a prospective, it is the intensive meaning in (ii) that best matches the interpretation of Santome *ka ka*.

In sum, *ka ka* and *ka ska* are closely related. In temporal clauses, these forms are synonymous, but there seem to be slight interpretation differences in imperatives and *ka ka* additionally insists on the action of the predicate that it modifies.

3.3.2.8. *Sa ska*

This progressive construction is very similar in meaning to the default progressive *sa ka~ska* and is particularly frequent in clauses that have temporal duration, as follows from (109-113). Like *ka ska* and *ka ka*, this construction has not been discussed in the relevant literature. Consider the following examples.

- (109) *Soku ola ku plaman sa ska da kodon, so ê lanta dôdôsu*
 then when KU morning be ASP give rope, then 3SG lift up RED-two
inen manu dê ni kama.
 3PL brother POS from bed

‘When morning was breaking, he woke up both his brothers.’

- (110) *Tudu inen mina-mina-mina fya se, ola san ka sa ska fla, non*
 all 3PL little-little-little leaveSP whenshe ASP be ASP speak 1PL
konsê.
 know

‘All these bits of leaves, when she is speaking about them, we know them.’

- (111) *Semple non sa ska kopla lôpa zêntxi.*
 always 1PL be ASP buy clothing people

‘We are always buying their clothes.’

- (112) *Inen dja se Zon sa ska kume ben.*
 3PL day SP Zon be ASP eat well

‘Lately, Zon has been eating well.’

- (113) *Modu ku kwa sa ska dêsé, sela non pidji mixikoji.*
 way KU thing be ASP go down must 1PL ask mercy

‘The way things are, we must ask for mercy.’

The temporal clause in (109) above shows that the progressive contains the perfect tense in the matrix clause due to its stative semantics. In fact, other constructions, for instance those in which temporal operators occur, show that there do not seem to be any meaningful differences between *sa ka~ska* and *sa ska* (114-115).

(114) *Djina onten, mwala sa ska dansa.*
since yesterday woman be ASP dance
'The woman has been/is dancing since yesterday.'

(115) *Mê dja, mwala sa ska dansa.*
'The woman is dancing at noon.'

The examples above show, however, that *sa ska*, more than *sa ka~ska*, privileges contexts where the progressive bridges an extended time span, even though it is compatible with a punctual time adverb such as *mê dja* 'at noon'. Note that the durative interpretation of *sa ska* often implies iteration or repetition as well.

From extensive testing with native speakers, I arrived at several conclusions regarding the use of *ska*, *sa ka* and *sa ska*. The first finding, as shown, is that these items are not in complementary distribution. However, are they syntactically and semantically in free distribution? This question is especially difficult to answer and I tend to believe that this particular area of Santome's grammar is subject to change and variation, although the fine details can only be ascertained by extensive fieldwork on several variables such as age and geographic location. In my transcriptions, I did not find any instances of *sa ska* in tape-recordings of speakers from the Batepá area but their overall use of non-contracted *sa ka*, which is otherwise quite uncommon, as follows from Table 7 on p. 118, was numerically more prominent than in other areas. Speakers from the other areas, on the other hand, produced a few instances of *sa ska* but produced a lower number of occurrences of *sa ka* than in Batepá speakers. I tentatively conclude that some speakers are more prone to use *sa ka* for the durative/iterative function, as an ongoing event from the past into the present, whereas others use *sa ska* preferentially or are perhaps shifting to this form. However, as it stands, this issue needs further field research.

While testing these markers, native speakers have repeatedly expressed the intuition that *sa ka* and especially *sa ska* show a tendency towards occurring in durative/iterative environments, although they also accept *ska* in these contexts. In the sense of Reichenbach (1947), the progressive markers *sa ka*, *ska~ska* and *sa ska* have two distinct interpretations. One whereby E(vent time) and S(peech time) converge, corresponding to the real present reading (is X-ing) and another whereby E(vent time) conflates with R(eference time) prior to S(peech time). The latter temporal construction

characterizes the English present perfect.⁶⁷ In the absence of sufficient historical data, it is impossible to reconstruct the evolution of these forms and predict their future functions. For instance, it might be the case that *sa ska* is specializing exclusively for the durative/iterative progressive function and *ska* for the real present alone, with a tendency for *sa ka* to become archaic.

The specializing scenario suggested depends very much on the origins of the *sa ska*-construction. From a diachronic point of view, there are two possibilities for arriving at the present construction, illustrated in (116a) and (116b):

- (116) a. [_{VP} *sa* [_{VP} *sa*] [_{AspP} *ka* [_{VP}]]]
 b. [_{VP} *sa* [_{AspP} *s(a)ka* [_{VP}]]]

The structure in (116a) sees *sa ska* as the historical result of a reduplicated verb form (*sa sa*), irrespectively of the syntactic details of reduplication, which would semantically match the fact that *sa ska* is especially prone to occur as a durative/iterative construction, since verb reduplication in Santome typically yields duration and/or iteration. Although I did not attest the construction *sa sa ka* in my spoken corpus nor in written texts, my informants accepted it and provided several examples, as for instance the following:

- (117) A *sa sa ka kôlê*.
 IMP be be ASP run
 ‘They are running and running.’

Since this construction is considered grammatical but atypical, for example compared with *ka ska~ka sa ka*, it might have disappeared over time once *sa ska* became grammaticalized.

The possibility sketched in (116b) suggests that *ska* entered this construction as a fully grammatical progressive marker that has started to replace *ka*. In this case, *sa ska* is presumably a recent change. In favour of this hypothesis is the fact that this construction is not attested in the literature on TMA in better described Ngola and

⁶⁷ Of course there are further nuances that go beyond the temporal interpretation of each construction. The English present perfect not only describes an ongoing event from the past, but also events that did take place in the past but bear current relevance at the time of speech.

Lung'ie, for these languages share a great many features in the TMA domain. Since *ska* is a functional progressive marker, the possibility that the copula is being reintroduced in analogy with the Portuguese progressive construction *estar a* (lit: 'be at') should also be considered. Yet, the analysis in (116b) fails to explain why speakers prefer to associate *sa ska* with duration and iteration.

3.3.2.9. *Tava ka ska, tava ka ka, tava sa ska*

It was mentioned earlier that the past imperfective *tava ka* incorporates both the past habitual and the past progressive and that there is no such construction as **tava ska*. However, Santome exhibits a past continuative/iterative progressive construction with *tava ka ska* or *tava ka ka*, as illustrated in (118-119), and pluperfect progressive, in (120).

- (118) *Mosu tava ka ska mintxila.*
 boy TNS ASP ASP lie
 'The boy was lying and lying.'
- (119) *Ke tava ka ka klêsê, ka klese...*
 house TNS ASP ASP grow ASP grow
 'The house was getting bigger and bigger.'
- (120) *Mosu tava sa ska fla von-von.*
 boy TNS be ASP speak nonsense.
 'The boy had been saying nonsense.'

Note that I didn't attest any of these structures in my corpus, but there is general agreement among consultants on its grammaticality.

3.3.2.10. Summary

The following table provides a summary of the grammatical functions discussed above.

Table 6. Semantic functions of tense-aspect markers (with dynamic predicates).

<i>Tense-aspect marker</i>	<i>Semantic interpretation</i>
<i>ka</i>	habitual, future
<i>sa ka~ska~xka</i>	present progressive, progressive present perfect
non-marked verb	perfect
<i>tava~ta</i>	pluperfect
<i>tava ka~ta ka</i>	past imperfective (past habitual / past progressive)
<i>ka sa ka~ka ska</i>	habitual/future progressive (iterative)
<i>ka ka</i>	habitual/future progressive (iterative); insistence
<i>sa ska</i>	present progressive, progressive present perfect
<i>tava ka ska</i>	past habitual progressive (iterative)
<i>tava ka ka</i>	past habitual progressive (iterative)
<i>tava sa ska</i>	pluperfect progressive

Table 7 below provides an indication of the relative proportions of the tense-aspect markers discussed in the previous sections, as they occur in my spoken data. Note that most of the spoken corpus consists of folk stories, so in other types of speech the proportions may be different.

Table 7. Occurrences of tense-aspect markers in the spoken corpus.

Tense-aspect markers	<i>Number of occurrences</i>
<i>ka + VP</i>	4,533
<i>ska~xka + VP</i>	1,181
<i>sa ka + VP</i>	27
<i>tava + VP</i>	14
<i>ta + VP</i>	25
<i>ta ta + VP</i>	1
<i>tava ka + VP</i>	36
<i>ta ka + VP</i>	34
<i>ka ska~ka xka + VP</i>	27
<i>ka sa ka + VP</i>	1
<i>ka ka + VP</i>	37
<i>sa ska + VP</i>	18
<i>sa sa ka + VP</i>	0
<i>tava ka ska + VP</i>	0
<i>tava ka ka + VP</i>	0
<i>tava sa ska</i>	0
total	5,934

As predicted, it can be concluded that, the habitual and the progressive marker are significantly prominent, not in the least because they accumulate several functions. I did not include the non-marked verb in the count because it would require a one by one count of all the instances of non-marked verbs. However, there is little doubt that Ø-marked verbs would compete with aspectual *ka* for the highest number of occurrences.

3.3.3. Stativity

One of the aspects of the tense-aspect system that has been poorly explored for Santome in particular, but was explored in a detailed fashion for Lung'ie and Ngola by Maurer (1995, 1997, forthc.), is the non-uniform behaviour of verbs with respect to the core tense-aspect markers. This difference is usually accounted for by means of a stative/non-stative opposition, a feature that came into the spotlight following the work of Bickerton (1975). The pairs of examples in (121) and (122) exemplify the temporal-

aspectual distinction between a stative predicate and a process with the bare verb and tense marker *tava*.

- (121) a. *Ê sêbê kwa se.*
3SG know thing SP
'He knows that.'
- b. *Ê tava sêbê kwa se.*
3SG TNS know thing SP
'He knew that.'
- (122) a. *Ê kôlê.*
'He ran.'
- b. *Ê tava kôlê.*
'He had run.'

The restrictions on stative verbs are often language-specific. It is well known that in English, for instance, stative verbs cannot occur in the progressive form *V-ing* (*I am knowing). Languages may exhibit a fine-grained typology of stativity, which is, for instance, the case of Russian, where a tripartite distinction can be made between permanent properties and relations, temporary states and permanent states (Spencer & Zaretskaya Ms.) according to syntactic and semantic tests.

Maurer (1995, 1997, *forthc.*) has argued that, in addition to dynamic verbs, Lung'ie and Ngola exhibit two types of statives that cannot be distinguished according to lexical aspect. On several occasions, he mentions, *en passant*, that this particularity also applies to Santome. The split between two types of stative verbs, he argues, follows from the morphosyntactic properties of these verbs. In other words, as in the examples above, TMA-markers overtly signal what type of verb class one is dealing with. The following table, based upon Maurer (1995, 1997, *forthc.*), considers the default semantic values associated to TMA-marking for each type.

Table 8. Functions of *ka* and \emptyset according to lexical aspect.

<i>Marker</i>	<i>States I</i>	<i>States II</i>	<i>Activities</i>
<i>ka</i>		present habitual present continuous	present habitual
\emptyset	present continuous	past perfective	past perfective
<i>tava ka</i>	n.a.	past imperfective	past imperfective
<i>tava</i>	past continuous	past-before-past	past-before-past

In this proposal, State I verbs take \emptyset to refer to present tense and *tava* to past tense, and lack the distinction between imperfective and perfective. State II verbs, on the other hand, take *ka* for present reference and have an aspectual distinction in the past: *tava ka* stands for past imperfective and \emptyset for past perfective. In other words, State I verbs typically lack event-internal structure. Comparing the lists of State I and State II verbs that Maurer provides for Ngola and Lung'ie, there appears to be a significant amount of overlap between these two languages. The following examples illustrate Maurer's proposal with my data from Santome.

- (123) *Ê kônsê tudu oso.* (State I)
 3SG know all bone
 'He knows all the bones.'
- (124) *Jina n kyê ni po, n ga mendu madêra.* (State II)
 since 1SG fall from tree 1SG ASP be afraid wood
 'Ever since I fell out of the tree, I'm afraid of wood.'
- (125) *Ê ka munja, ê ka pya, ê ka kôlê wê.* (Activities)
 3SG ASP stop 3SG ASP look 3SG ASP run eye
 'He stops, he gazes, he looks around.'

The proposed difference between a State I verb and a State II verb can be further illustrated in examples (126-127), where *gôgô* 'to enjoy, to love, to like' contrasts with *kontlê* 'to hate'.⁶⁸

⁶⁸ Note that in Russian, 'to love' is a permanent state and 'to hate' a temporary state (Spencer & Zaretskaya Ms.).

(126) Ê *gôgô* ku fesa.

3SG like with party

‘He likes parties.’

(127) Ê *ka kontlê* fesa.

3SG ASP hate party

‘He hates parties.’

Morphosyntactically, the verb *kontlê* behaves in all respects like a dynamic verb, despite the fact that it describes a permanent property of the subject. Other verbs that exhibit the same properties as *kontlê* are, for instance, *mendu* ‘to be afraid of’ and *ta* ‘to live, to be at’. That lexical aspect is indeed an insufficient criterion to distinguish between verb classes follows from the following example with the epistemic verbs *kunda* ‘to think’ and *pensa* ‘to think’, which have approximately the same meaning.

(128) Ê *kunda* kuma omali sa zulu.

3SG think that sea be blue

‘He thinks the sea is blue.’

(129) *Inen pensa* kwa zo fe.

3PL think thing then do

‘They thought before doing it.’

A special case that has been highlighted in Maurer’s work on Lung’ie and Ngola is the verb *pô* ‘can, may, be able to’, which exhibits two distinct patterns. This also applies to Santome, where the bare verb expresses a deontic reading, in (130), and the verb preceded by *ka* an epistemic reading, in (131).

(130) *Bô pô* bixi mina bô lôpa.

2SG can dress infant POS clothing

‘You may dress your infant.’

(131) *Kuma n ga pô* fla ku ngê sê sêbê ken sa sun?

how 1SG ASP can speak with people without know who is man

‘How can I speak to someone without knowing him?’

In sum, the examples show that lexical aspect does not yield systematic syntactic patterns in the GGC. Moreover, the differences between State II verbs and dynamic verbs in Maurer’s proposal are related to lexical aspect and not to syntax, since both types exhibit the same pattern of tense-aspect marking. In the remainder of this section I will therefore outline a syntactic approach to the verb classes.

In the light of the discussion above, my first assumption is that Santome exhibits a primary split between [+States] and [-States]. In order to establish this distinction, I will apply Maurer’s diagnostic tool for the identification of State I verbs, whereby a state can be diagnosed if the bare verb in a declarative main clause receives a present (continuous) interpretation and if it has a past (continuous) interpretation when the verb takes tense marker *tava~ta*, as illustrated here in (132):

- (132) a. *Ê sêbê tudu inen kwa se.*
 3SG know all 3PL thing SP
 ‘He knows all these things.’
 b. *Ê ta sêbê tudu inen kwa se.*
 3SG TNS know all 3PL thing SP
 ‘He knew all these things.’

The following verbs behave in this manner: *sen* ‘to exist’, *tê* ‘to have’, *sa* ‘to be’, *kônsê* ‘to know, to recognize’, *mêsê* ‘to want, to love, to desire’, *pô* ‘can, may’ (permission), *gôgô* ‘to like, to enjoy’, *ngosta~gosta* ‘to like, to enjoy’, *fata* ‘to lack, to miss’, *kia* ‘to want’ and *kunda* ‘to think’.

I will now argue that the specific behavior of stative verbs with respect to tense-aspect marking, to be shown next, consists of a distinction between stage level and individual level predicates. In order to support this claim, consider the following contrast between the use of the verb *ngosta~gosta* ‘to enjoy, to love, to like’ with and without aspect marker *ka*, as in (133) and (132) respectively:

- (133) a. *Mosu se ka ngosta d’uswa muntu.*
 boy SP ASP like of sour palm wine much
 ‘The boy in question loves sour palm wine a lot.’

- b. *Makaku ka gosta bôbô.*
 monkey ASP like ripe banana
 ‘Monkeys like ripe bananas.’
- c. *N sa ome ku ka gosta mwala muntu.*
 1SG be man that ASP like woman much
 ‘I’m a man who likes women a lot.’
- (134) a. *Mina se gosta d’ami muntu.*
 girl SP like of-1SG much
 ‘The girl in question likes me a lot.’
- b. *N na ngosta kala dê fa.*
 1SG NEG like face POS NEG
 ‘I don’t like the way he looks’ (i.e. I’m worried about his aspect)
- c. *Maji n na gosta di klonveson se kompa mu fô.*
 But 1SG NEG like of conversationSP godfather POS NEG-EMPH
 ‘But I don’t like the way my godfather talks.’

The verb *ngosta~gosta* is included in the list of syntactically stative verbs, something which is expected from its lexical aspect. In Maurer’s classification, *ngosta~gosta* would be a problem-case because it exhibits a present reading both with the \emptyset -marker and with *ka*, and would thus be a member of both the State I and the State II classes. I will argue, however, that *ngosta~gosta* is a [+State] verb and that the difference between the sentences in (133) and in (134) is related to the internal structure of the predicate, i.e. whether we are dealing with a stage level predicate or an individual level predicate.

The distinctio between stage level and individual level goes back to Milsark (1974) and has ever since been explored in syntax and semantics by a number of authors (e.g. Carlson 1977, Diesing 1992, Kratzer 1995). In a nutshell, stage level predicates typically express temporary properties and therefore predicate over stages, whereas individual level predicates are concerned with permanent properties and predicate over individuals.

The difference between the situation noted here and the usual assumptions about stage levels is that the sentences in (133) denote not a single stage level event but rather a recurring sequence of stages that yield a habitual state reading. The sentences in (134),

on the other hand, have a permanent character and comply with the definition of individual level predicates. In order to formalize this distinction, I will adopt a suggestion made by Manninen (2002) who, following earlier findings by Carlson (1977), proposes that individual-level predicates can be divided into:

- (i) “*habitual* predicates, which express generalisations over a large number of properties which are characteristic of an individual over an extended period of time.”
- (ii) “*property* predicates, which describe properties which are characteristic of an individual over an extended period of time.”

Thus, when a stative verb like *ngosta* is preceded by habitual aspect marker *ka*, it means that the subject is repeatedly engaged in a certain activity (for example: drinking palm wine, eating bananas, liking women) and that the truth value of this habit holds irrespectively of the time interval between each iterated event of which the habit is comprised. The use of the bare verb, without the aspect marker, on the other hand, signals a continuous state that cannot be decomposed in subevents. The distinction between stage and individual level predicates is confirmed by adding a temporal adjunct PP, *ni dja djingu* ‘on Sunday(s)’, to the stative predicate.

- (135) *Mina se gosta d’ami muntu ni dja djingu.*
 girl SP like of-1SG much on day Sunday
 a. ‘The girl liked me very much on Sunday(s).’
 b. * ‘The girl likes me very much on Sunday.’
 c. * ‘The girl likes me very much on Sundays.’

- (136) *Mina se ka gosta d’ami muntu ni dja djingu.*
 girl SP ASP like of-1SG much on day Sunday
 ‘The girl likes me a lot on Sundays.’

In example (135), adding a temporal point blocks the present continuous reading in (135b) and obligatorily yields a perfective reading as in the interpretation in (135a). Interpretation (135c) is also precluded because the plural interpretation of the PP transforms the sentence into a habit. But to derive a habitual state, the use of preverbal

ka is required, as follows from the reading associated with (136), which can only be thus interpreted if the PP receives a plural reading.

Stage level interpretations are also obtained when a construction contains or implies several stages, as illustrated in (137a-b):

- (137) a. *Ũa dja n ga gôgô ku pixi; ôtlô dja n ga gôgô ku kani.*
 one day 1SG ASP like with fish other day 1SG ASP like with meat
 ‘One day I like fish; the other day I like meat.’
- b. *Ũa dja n ga gôgô ku Zon; ôtlô dja n ga gôgô ku Zose.*
 ‘One day I like Zon; the other day I like Zose.’

Examples (137a-b) can be described as a habitual transition between two stages. In (137a), the occurrence of the aspect marker is fully expected if we assume that liking fish or meat is the sum of a number of atomic events that brings about a habit. Liking Zon or Zose, in (137b), however, is a property predicate but, unexpectedly, both clauses are aspectually marked (by *ga*). In these cases, I assume that the change of state triggers a stage level interpretation. Even with a strongly stative verb such as existential *sen* ‘to be, to exist’, the transition between states can be marked with *ka*, as in (138), although in this case native speakers do not sense a significant difference between the sentence with *ka* and the one without it.

- (138) a. *Ũa dja konxensa sen; ôtlô dja konxensa na sen fa.*
 One day conscience exist other day conscienceNEG exist NEG
 ‘One day conscience exists; the other day it doesn’t.’
- b. *Ũa dja konxensa ka sen, ôtlô dja konxensa na ka sen fa.*
 ‘One day conscience (usually) exists; the other day it doesn’t.’

Furthermore, the following examples show that the aspect marker is also required when the transition involves a past or a present state, in (139), instead of a recurring transition as in (137) and (138) above, and it is also required when the stative predicate is temporally anchored, as exemplified in (140).

- (139) *Noxtempu n tava ka gôgô ku Zon, maji djina 1970 n ga gôgô*
 formerly 1SG TNS ASP like with Zon but since 1970 1SG ASP like
ku Zose.
 with Zose
 ‘Formerly I liked Zon, but since 1970 I like Zose.’
- (140) *Ola n tava mosu, n tava ka gôgô ku Maya.*
 when 1SG TNS boy 1SG TNS ASP like with Maya
 ‘When I was a boy, I liked Maya.’

Does this hypothesis hold for other stative predicates as well? Let us consider a typical stative verb like *tê* ‘to have, to possess’. In the examples in (141), the verb takes *ka* whereas in the examples in (142) it does not.

- (141) a. *Soku ola sungê ka sa novu, ê ka tê migu, ê ka tê*
 so whenyou ASP be young 3SG ASP have friend 3SG ASP have
yô ngê dê.
 many person POS
 ‘So when you are young, you have friends, you have a lot of people (around you).’
- b. *Sun se sa ngê ku ka sêbê kwa tela.*
 man SP be person REL ASP know thing country
 ‘The man in question used to know things about the country.’
- (142) a. *Non na tê men fâ, non na tê pe fâ.*
 1PL NEG have mother NEG 1PL NEG have father NEG
 ‘We don’t have a father and mother.’
- b. *N sêbê kwa n ga ba fe.*
 1SG know thing 1SG ASP go do
 ‘I know what I will do.’

The verb *tê* is a typical stative in many languages and the present reading in the absence of any tense-aspect marker in (142) confirms that this is also the case in Santome. Maurer classifies *tê* as a State I in Ngola and Lung’ie. The examples in (142a-b) denote

a property predicate that lacks internal eventive structure, i.e. the truth value of the predicate holds at any time interval. Example (141a), on the other hand, is a stage-level predicate with a temporal clause delimiting the stage (being young, youth) for which the affirmation holds, with the implication that it only holds for that specific time slice but not beyond. Therefore, the verbs *sa* and *tê* are aspectually marked by *ka*. Although an individual level predicate, example (141b) denotes a recurring habit and not an intrinsic property of the subject. Note further that in the presence of an operator that locates a state in the future, for instance a temporal adverb, *ka* receives a future reading. Compare the following sentences:

- (143) *Ê sêbê kwa se.*
 3SG know thing SP
 ‘He knows this.’
- (144) *Amanhan, ê sêbê kwa se.*
 ‘Tomorrow, he knows this.’
- (145) *Amanhan, ê ka sêbê kwa se.*
 ‘Tomorrow, he will know this.’

In the light of these examples, which show a property that also applies to dynamic predicates, I assume that *ka* is not inherently specified for the value [Future] and I therefore treat this marker as an aspectual marker in all the relevant contexts.

Summing up the discussion on stative predicates, it followed that they can be divided into stage level predicates and individual level predicates. The latter can be further divided in properties and habitual states. The following table summarizes the findings above.

Table 9. Types of stativity and tense-aspect marking.

<i>States</i>			
	individual level		stage-level
	properties	habitual states	
present	\emptyset	<i>ka</i>	<i>ka</i>
past	<i>tava</i>	<i>tava ka</i>	<i>tava ka</i>

The table shows that individual level properties can be formally distinguished from habitual states and stage level predicates, whose behaviour is similar to the behaviour of dynamic verbs. The impossibility of aspect marking in individual level property predicates matches the consensual assumption that these predicates lack internal structure. I therefore propose that the difference between individual level properties, on the one hand, and all other predicate types, on the other hand, including non-property states, has a syntactic counterpart. The former predicates are exceptional, I claim, in lacking an aspectual projection in their structure, a proposal that has been made by several authors for stative predicates in general (e.g. MacDonald 2006). I will further discuss this hypothesis in section 3.3.6.

Note further that in most cases an eventive reading of states can be coerced. In example (146-147), the bare verb, *sêbê* and *kônsê*, are located in the past by adverb *onten* ‘yesterday’ and therefore get an inchoative perfective (past perfect) reading. In (148-149), the verb is preceded by progressive aspect marker *ska* and therefore the predicate receives a non-permanent interpretation.

(146) *Onten, ê sêbê kuma suba sôbê ni Lixboa.*

Yesterday 3SG know that rain rain in Lisbon
 ‘Yesterday, he came to know that it rained in Lisbon.’

(147) *Onten, ê kônsê tudu ngê.*

yesterday, 3SG know all person
 ‘Yesterday, he met everybody.’

(148) *Pema na ska mêsê to bwadu fa.*

palm tree NEG ASP want drip good-PP NEG
 ‘The palm tree doesn’t want to drip well.’⁶⁹

(149) *Ôtlô ku na ska tê kloson fa.*

Other REL NEG ASP have heart NEG
 ‘Another one who is insensitive.’

Therefore, I assume that in these cases the aspectual structure of the predicates has to be projected in syntax, contrasting, once again, with the treatment of property predicates proposed above.

⁶⁹ In other words, the palm juice wasn’t dripping well out of the tree into the recipient.

3.3.4. Relative tense

Comrie establishes a difference between languages exhibiting absolute tense and those exhibiting relative tense. Absolute tense implies that temporal reference is established with respect to the present moment, whereas relative tense may have other reference points that are not the present moment, for instance, some previous time reference anchored in discourse. According to this definition of tense, it can be shown that Santome behaves like a relative tense language, since the time reference of events is frequently determined contextually and not with respect to the time of utterance (Maurer 1997, Schang 2000). Narrative texts such as folk tales clearly underscore this claim. Consider the following passage in a folk story.

(150) *Ê ba awa ba laba platu. Ê laba platu. So, ola ê ska*
3SG go river go wash plate 3SG wash plate then when 3SG ASP

luma platu, ê fat'e ãa platu. Kuma awa tava ka dêsê
store plate 3SG lack-3SG one plate Since river TNS ASP descend

ku fosa, platu dêsê n'awa. So, ê ska sola. Ê ska sola,
with strength plate descend in-river then 3SG ASP cry 3SG ASP cry

ska dêsê n'awa, ska glita: kuma n ga ba ke
ASP descend in-river ASP scream how 1SG ASP go house

ba fada mama mu kuma platu mama plêdê?
go tell mother POS that plate mother loose

‘She went to the river to wash plates. And so she did. Then, when she was storing the plates, one plate was missing. Since the river had a strong current, the plate was taken by it. Then, she started crying. She was crying and went into the river, screaming: How can I go home and tell my mom that I have lost her plate?’

Although this passage doesn't correspond to the start of the story, the first two sentences frame it in the past, since bare dynamic verbs typically have a perfective interpretation. The use of progressive *ska* in the third sentence, however, contrasts with the preceding sentences in the sense that *ska* typically conveys a progressive present meaning, for instance, in a sentence out of the blue. The past imperfective, which corresponds to the interpretation of *ska sola*, would normally be expressed by *tava ka* (cf. *tava ka dêsê*) or *ta ka*. Therefore, it can be concluded that temporal reference of the progressive *ska* in the third sentence and in the remainder of the passage is anchored by the perfect. It follows that discourse (e.g. Givón 1982, Michaelis 1993) plays an important role with respect to framing tense in this language.

The following examples of a temporal clause and two complementives introduced by finite complementizer *kuma* show that relative tense also operates at a more local level

(151) *Ola se karu na ta sen fa, a ka dêsê n'ope.*
 when SP car NEG TNS exist NEG IMP ASP go down by foot
 'And when there was no car, one went down on foot.'

(152) *N konta kuma ê ska bi.*
 1SG count that 3SG ASP come
 'I expected that he would come.'

(153) *Ê tava sêbê kuma mwala ba ke.*
 3SG TNS know that woman go house
 'He knew that the woman had gone home.'

In these examples it is the first clause that establishes the temporal reference of the following clause. All the examples have in common that an aspectual marker is bound or controlled by tense outside its clausal domain. Since there are no reasons to believe that the tense-dependent clauses above do not project TP in their structure, I assume that tense can be inherited non-locally. Although this usually implies a relation whereby the tensed domain linearly precedes the tense-dependent domain, strict linear precedence is not necessarily required in order to establish a tense-chain, as follows from inverting the clauses of example (151).

(154) *A ka dêsê n'ope, ola se karu na ta sen fa.*
 'One went down by foot when there was no car.'

3.3.5. Projecting aspect

This section focuses on the properties of core aspectual marking and the representation of aspect in syntactic structure. Compared to tense and its representation as the functional projection TP in generative grammar, the treatment of aspect has been quite more heterogeneous in the specialized literature. Many authors have proposed one or more aspectual projections between VP and TP, but the label of this projection varies significantly. Labels such as AspP, TransitivityP, EventP, AuxP, vP or AgroP all fit within this tradition.⁷⁰ In part, the observed incongruence between the label for tense and that for aspect derives from the fact that sentences often exhibit more than one type of aspectual information. Given its transparency, I will adopt the label AspP (e.g. Gonçalves, 1996, Stowell 1993, Thompson 1996, Baker 1997, Borer 1997, Cinque 1999, Matsuo 2001).

Given the general tendency to refine clause structure over the past decades (e.g. Pollock 1989, Rizzi 1997, Cinque 1999), how much functional structure should be projected in syntax, language-internally and cross-linguistically, and how it is projected have become major issues. The lack of consensus on aspect encoding leads to strongly contrasting analyses. In the functional cartography proposed by Cinque (1999), several AspPs can be instantiated in clause structure, a proposal that has been adopted for individual languages such as Gungbe (Aboh 2004) and Jamaican Creole (Durrleman 2000). Recently, it has also been proposed that aspectual markers and other functional material, including negation, are recursively adjoined heads under TP (Costa & Pratas 2004, Pratas 2004). An important distinction between these two proposals is thus not so much about what should project in syntax but rather about where it should project and what features the projections comprise. In addition, it becomes crucial to discuss the categorial features of the TMA-material in order to determine whether they are heads or whether they have phrasal status or whether these items are truly functional or exhibit verb or auxiliary features. These distinctions are of course not always easy to make. Below, I will seek to address these issues.

In the literature on Santome, especially Ferraz (1979), it is habitually assumed that core aspect comprises two aspect markers, namely habitual/future *ka* and

⁷⁰ In a certain sense, the absence of a syntactically uniform hypothesis for the VP-modifying domain is a reflex of the complex state of research on this module within semantics (e.g. Sasse 2002).

progressive *sa ka~ska~xka*. This bipartition does not carry over to past imperfectives, where *tava ka* or its free variant *ta ka* can be interpreted as a past progressive or a past habitual, whereas the virtual construction **tava ska* is lacking and, to the best of my knowledge, is not attested in synchronic or diachronic sources.

First, it is necessary to determine the syntactic status of the aspect markers above. Consider the following evidence from participle constructions (155), VP-ellipsis (156), question-answer pairs (157), VP-fronting (158) and adverb placement (159).

(155) **Kinte ka balidu.*

garden ASP swept

(156) **Zon ka [bali kinte]_i, maji Maya na ka [-]_i fa.*

Zon ASP sweep garden but Maya NEG ASP NEG

(157) Q: *Zon ka bali kinte?*

Zon ASP sweep garden

‘Does Zon sweep the garden?’

A: *Efan, ê ka *(bali).*

yes 3SG ASP sweep

‘Yes, he does.’

(158) a. *Bô ka bali kinte.*

2SG ASP sweep garden

‘Zon sweeps the garden.’

b. *Bali kinte so bô ka *(bali).*

sweep garden FOC 2SG ASP sweep

‘Sweep the garden is what he does.’

(159) *Zon ka (*adverb) bali kinte.*

Zon ASP adverb sweep garden

‘Zon swept the garden / sweeps gardens.’

These tests show conclusively that:

- (i) *ka* has functional properties;
- (ii) *ka* behaves like a bound morpheme with respect to the verb.⁷¹

⁷¹ It is also illustrative that, in written texts in Santome, *ka* and the verb are often represented as a single, contracted form.

Therefore, I assume that *ka* heads a lexicalized functional projection labeled AspP. It was shown in section 3.3.2.1 that the functions of *ka* are not restricted to habitual aspect, which, however, is its most common interpretation. Instead of storing several different *ka*'s in the lexicon, I propose that this imperfective marker is semantically underspecified, acquiring its features compositionally in the derivation. The counterpart of this marker is the perfective zero marker or non-marked verb. I assume that AspP always projects with dynamic predicates, even when an imperfective predicate lacks a lexical counterpart, as can be observed in the following aspectual construction.

- (160) a. *Punda n sa kwaji Ø be mu za.*
 because 1SG be about to go PSR already
 'Because I'm about to leave.'
 b. **Punda n sa be mu za.*
- (161) a. *Nen sa kwaji Ø bila ve za, so inen pali dôsu mina.*
 3PL be about to turn old already then 3PL give birth two child
 'They were already getting old, so they had two children.'
 b. **Nen sa bila ve za, so inen pali dôsu mina.*

The absence of *ka* in (160a) and (161a) cannot be due to the properties of *sa*, a fact which follows from the ungrammaticality of the b. sentences without *kwaji*. Note also that the aspect marker *ka* can be realized overtly in these constructions, as shown in the following example:

- (162) *Nen sa kwaji-kwaji ka xiga ke.*
 3PL be almost-almost ASP arrive house
 'They were about to arrive home.'

In the light of this evidence, I assume that the aspectual adverb *kwaji* is base-generated as a left-adjunct to AspP signaling the presence of this functional projection. To be sure, there is evidence that *kwaji* is adjoined to AspP and not a modifier of *sa*. Consider the following perfective example:

- (163) *Ê kwaji Ø kume tudaxi.*
 3SG almost eat everything

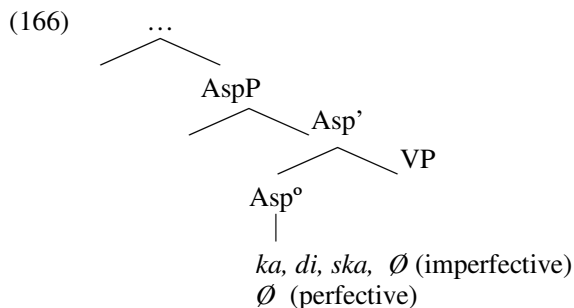
‘He almost ate everything.’

In this case, I assume that *kwaji* is a left-adjunct of AspP in a similar fashion, with the only specificity that the head of AspP is realized as the perfective zero marker. Finally, the head of AspP can also be filled by other lexical items, namely *di* and *ska*.

(164) *Ome mu sa kwaji di bi.*
 Man POS be almost ASP come
 ‘My husband is about to arrive.’

(165) *Non sa kwaji ska klonvesa.*
 1PL be almost ASP converse
 ‘We are almost conversating.’

These structures exhibit the same properties as the *sa ka-ska*-construction, which will be discussed in more detail below. A preliminary conclusion is that AspP is a standard projection in Santome’s basic clause architecture, as represented in (166), where the head of AspP can be filled by *ka*, *di*, *ska* and \emptyset for imperfective readings and by \emptyset for perfective readings.



The apparent ambiguity between the zero marker for perfective and imperfective can be solved by assuming that *sa*, albeit not a functional aspect marker, has properties that instantiate imperfectivity in the relevant construction.⁷²

I will now consider the properties of the present progressive/perfect construction. According to Ferraz (1979: 82), *ska* and *xka* are contracted forms of *sa ka*.

⁷² Note also that there is nothing special about the fact that \emptyset can be assigned several functions. The same happens, for instance, in the case of Bare Noun Phrases, which in Santome can be [+/-definite] and [+/-singular] depending on grammatical and extra-grammatical properties (cf. Alexandre & Hagemeijer, forthc.)

Native speakers confirm that these forms are indeed interchangeable. Since I have already analyzed the syntactic properties of *ka*, I will first apply the same tests used for *ka* to *sa* in this construction. Note, however, that *sa*, as mentioned above, cannot immediately precede the verb. Therefore, in (167-170), I test the *sa ka*-construction for the same properties that were evaluated in examples (155-159).

(167) *Kinte sa (*ka) balidu.*

garden be (ASP) swept

(168) *?Zon sa ka [bali kinte]_i, maji Maya na sa (*ka) [-]_i fa.*

Zon be ASP sweep garden but Maya NEG be (ASP) NEG

(169) Q: *Zon sa ka bali kinte?*

Zon be ASP sweep garden

‘Does Zon sweep the garden?’

A: *?Efan, ê sa (*ka).*

yes 3SG be ASP sweep

‘Yes, he does.’

(170) a. *Bô sa ka bali kinte.*

2SG be ASP sweep garden

‘Zon sweeps the garden.’

b. **Bali kinte so bô sa *(ka bali).*

sweep garden FOC 2SG be (ASP)sweep

‘Sweeping the garden is what he is doing.’

The results show a clear contrast with *ka*, since *sa* responds positively to several of the tests, despite some variation in judgment among native speakers with respect to the acceptability of both ellipsis, in (168), and isolation in question-answer pairs, in (169). In these cases, the preferred strategies are clearly stripping or repeating the main verb. Predicate focus in (170b) is ungrammatical. This means that *sa* does not exhibit the fully functional behavior of the aspect marker *ka* but also lacks the properties typically assigned to full-fledged verbs.

As for adverb placement, it was already shown that *kwaji* can intervene between *sa* and *ka*. As illustrated below, I have tested other adverbs in this position. Note that, to some extent, there is variation of the speaker’s judgments. The sentence final position is always considered the best option for the adverbs in question.

- (171) a. *Ê sa (?!?!ten) ka kume (ten).*
 3SG be (also) ASP eat (also)
 ‘He is eating as well.’
- (172) a. *Ê sa (?!/*tan) ka kume (tan).*
 3SG be (only) ASP eat (only)
 ‘He is only eating.’
- (173) a. *Zon sa (?!?!* pikina-pikina) ka kume kwa se (pikina-pikina).*
 Zon be little by little ASP eat thing SP (little by little)
 ‘Zon is little by little eating this.’
- (174) a. *Zon sa (?!?!leve-leve) ka kume kwa se (leve-leve).*
 Zon be (calmly) ASP eat thing SP (calmly)
 ‘Zon is calmly eating this.’

The grammaticality judgments show that there is no consensus among native speakers regarding the status of these sentences. Since *pikina-pikina* and *leve-leve* can be considered aspectual adverbs like *kwaji*, the conditioned acceptability or even ungrammaticality of (173) and (174) is hard to relate to a semantic constraint. Moreover, the hesitation with focus adverb *ten* and *tan* in (171) and (172) suggests that the prosodic weight of these adverbs – and hereto related a possible distinction between X° and XP adverbs - cannot be the cause of hesitation.

Finally, the present imperfective construction can be preceded by aspect marker *ka* (175):

- (175) *Ê ka {sa ka~ska~xka} da vin se novu-novu.*
 3SG ASP ASP give wine SP young-young
 ‘It [palm tree] is usually giving very young wine.’

Note that tense marking on *sa* (**tava sa*) is ungrammatical, but this follows from the fact that *tava* and *sa* form a suppletive pair. It goes without saying that the constructions involving *sa ka*, *sa di* and *sa ska* also behave as monoclausal domains with respect to tense and negation. Thus, it can also be concluded that *sa* is depleted of lexical meaning and this results in syntactic defectivity. This conclusion underlies my assumption that *sa*

is an auxiliary verb (Aux). The lack of consensus among native speakers concerning some of the structures with *sa* above underlines the fact that the aspectual construction has to some extent, and presumably to different extents, grammaticalized. I am adopting the term “auxiliary” only as a descriptive means to cover those items that are not fully functional nor fully lexical.⁷³

A remaining question is whether the contracted variants of *sa ka*, namely *ska~xka*, show the same behavior as the non-contracted form. It follows from (176-179) that this is not the case.

(176) **Kinte ska balidu*.

garden ASP swept

(177) **Zon ska [bali kinte]_i, maji Maya na ska *(bali) [-]_i fa*.

Zon ASP sweep garden but Maya NEG ASP (sweep) NEG

(178) Q: *Zon ska bali kinte?*

‘Does Zon sweep the garden?’

A: *Efan, ê ska *(bali)*.

‘Yes, he does.’

(179) a. *Bô ska bali kinte*.

2SG ASP sweep garden

‘Zon sweeps the garden.’

b. **Bali kinte so bô ska *(bali)*.

sweep garden FOC 2SG ASP sweep

‘Sweeping the garden is what he is doing.’

These examples show at least that the contraction of *sa* and *ka* did not yield a new verb with the meaning ‘to be at’, otherwise the sentences above would presumably be grammatical. Therefore, I argue, the aspect marker *ka* is still identified as a functional head and triggers the ungrammaticality of the examples above. In other words, it is *sa* that started cliticizing phonologically onto the aspectual head *ka*, and not the functional head that was incorporated into the verb *sa*, which also helps to explain why *sa* responds less well to the tests for verbhood above.

⁷³ For instance, Gonçalves (1996) shows that in European Portuguese auxiliaries *latu sensu* often behave quite differently with respect to syntactic tests.

The *sa ska*-construction confirms these developments within Santome's aspectual complex. This construction, described in section 3.3.2.8, has the meaning of an ongoing progressive.

- (180) *Inen dja se Zon sa ska kume ben.*
 3PL day SP Zon be ASP eat well
 'Zon is eating well these days.'
 'Zon has eaten well these days.'
 'Zon has been eating well these days.'

If we look at adverb placement of *kwaji* in this construction, as compared to other progressive constructions, it follows that there are good arguments to treat *ska* as a functional head. Note that *kwaji* can readily intervene between *sa* and *ka* in (181a) and (182a), but most informants have serious doubts about *kwaji* preceding *ska* in (181b) and consider this ungrammatical altogether for (182b). This sharply contrasts with the *sa ska*-construction, in (183), where the adverb is able to precede *ska*.

- (181) a. *Ê sa kwaji ka xiga ke.*
 3SG be almost ASP arrive house
 'He's almost arriving home.'
 b. *?/??/*Ê kwaji ska xiga ke.*
 (182) a. *Ê ka sa kwaji ka xiga ke.*
 b. **Ê ka kwaji ska xiga ke.*
 (183) *Ê sa kwaji ska xiga ke.*
 3SG be almost ASP arrive house
 'He's almost arriving home.'

I assume that this adverb is only able to precede *ska* in (183) because *ska* has fully grammaticalized as a functional progressive marker in this construction, heading AspP. All in all, I believe that the hesitation with respect to the syntactic tests applied to *sa ka* and the evidence of adverb placement, for instance *kwaji*, are best explained in a scenario of language change and variation. Even though for most speakers *ska* appears to be a full-fledged functional aspect marker, upon inquiry they do not hesitate to consider it the contracted form of *sa ka*, whose frequency in speech is low. Yet, my

informants are reluctant to accept a sequence like *kwaji ska*. Other speakers seem to use *sa ka* more actively, although still in much lower proportions than *ska*. The *sa ska*-construction, albeit not very productive, corroborates the fully functionalized status of *ska* in some constructions. However, that *ska* hasn't penetrated all progressive constructions as a functional item can be concluded from the fact that Santome lacks the **tava ska*-construction, which would virtually be the past progressive counterpart of *sa ska*.

At first glance, it may then look as if the progressive *sa ka*-construction is on its way to fully grammaticalizing into its contracted form *ska~xka*. This intuition seems to be confirmed by the number of occurrences in my spoken corpus, as shown in Table 10.

Table 10. Number of occurrences of the progressive in the spoken corpus.

	number of occurrences	% (approx.)
<i>sa ka</i>	90	6
<i>ska~xka</i>	1371	94
Total	1461	100

Nevertheless, these numbers in support of a possible grammaticalization of *sa ka* should be treated with some caution. Despite being much less extensive than my spoken corpus, the two older written sources that are available for Santome, namely Negreiros' (1895) chapter and pamphlets written by Bonfim in the 1920s and 1950s, show not only that *sa ka* and *ska* already existed at the time of their publications but suggest that *ska* was already the strongly predominant form.

Table 11. Number of occurrences of the progressive in older sources.⁷⁴

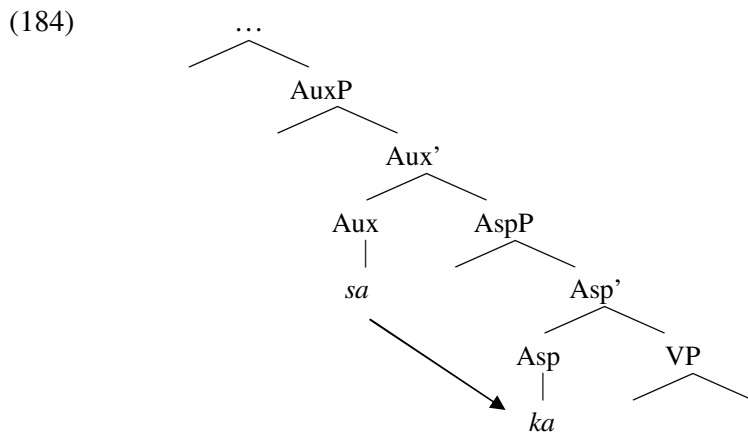
	<i>Negreiros</i>		<i>Bonfim</i>	
	tokens	%	tokens	% (approx)
<i>sa ka</i>	1	92	1	1,50
<i>ska</i>	12	8	65	98,50
Total	13	100	66	100

⁷⁴ Where necessary, I have adapted the original orthography and counted a number of repetitions in repeated lines in a poem in Negreiros (1895: 347) as a single instance.

Although it is unclear to what extent these written data can be considered representative of spoken Santome during the respective historical stages, the use of both the contracted and the non-contracted form show that both progressive constructions were known to these authors. Moreover, if their writings were to be influenced by the Portuguese progressive construction (*estar a*), one would definitely expect more instances of the non-contracted form, yet such instances are not found. Therefore, there are no obvious signs of grammaticalization of the progressive construction in the past century. Instead, the variation appears to be rather stable, to judge from the available sources.

What can still be concluded, however, is that the abovementioned restrictions on the *sa ka*-construction (ellipsis, adverb placement) show that this non-phonologically reduced construction is itself perceived as a more or less grammatical chunk by native speakers, which explains why the grammaticality judgments are considerably variable from speaker to speaker. Note further that I did not attest in older sources any constructions with, for instance, *sa ska* and *ka ska*, but this might simply be a gap in the data, given the low number of occurrences of these constructions in the contemporary language (cf. Table 10).

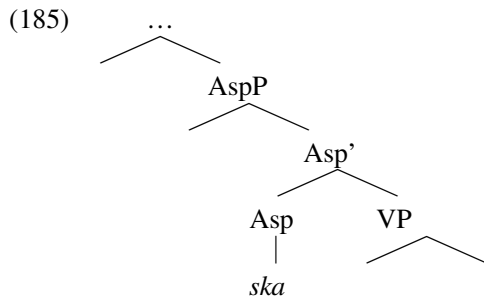
Considering the variation between *ska* and *sa ka* and the results of the syntactic tests, I assume that the reduced form *ska* is diachronically *or* synchronically the outcome of a post-syntactic phonological process whereby *sa* and *ka* merge(d). The synchronic situation is illustrated in tree (184).



If the necessary syntactic conditions are met, i.e. the absence of intervening material (e.g. *kwaji*) *sa ka* may post-syntactically contract into its short forms *ska~xka*, in the direction indicated by the arrow. In the spirit of Marantz (1988) and Halle & Marantz

(1993), this is a typical case of merger, whereby two adjacent heads merge into one in the phonological component but each heads remains visible for syntax. This tree structure presumably represents how a minority of speakers still deals with *ska* at a synchronic level.

It follows that for most speakers *ska* is already stored in the lexicon as a functional item that provides a progressive reading, i.e. *ska* is not visible for syntax anymore. For these speakers, the *sa ska*-construction and all other instances of *ska* are lumped together and receive the same analysis. Following again Halle & Marantz (1993), this would be a case of fusion and not of merger. As a consequence, the syntactic representation is different from the tree in (182), with *ska* heading AspP:



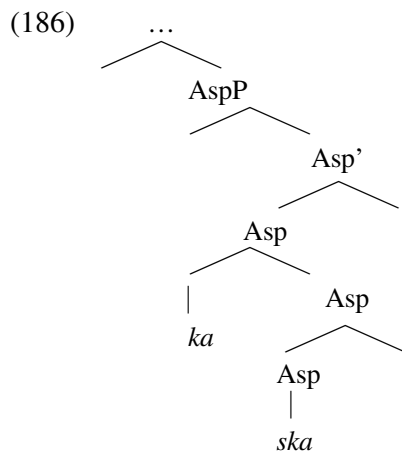
In this construction, *ska* is analyzed as a functional head with a status similar to *ka*. In fact, from a broader point of view, it can be concluded that as a consequence of phonological cliticization, followed by syntactic restructuring, *ska* is arguably specializing in the progressive function of *ka*. Thus, there is arguably a diachronic change taking place towards the tree structure in (185).

Analyzing *ska* as a full-fledged progressive marker raises the question of how to treat *ka ska*, where two functional items cluster together. Assuming that both items are heads, in agreement with the discussion in this section, there are several solutions:

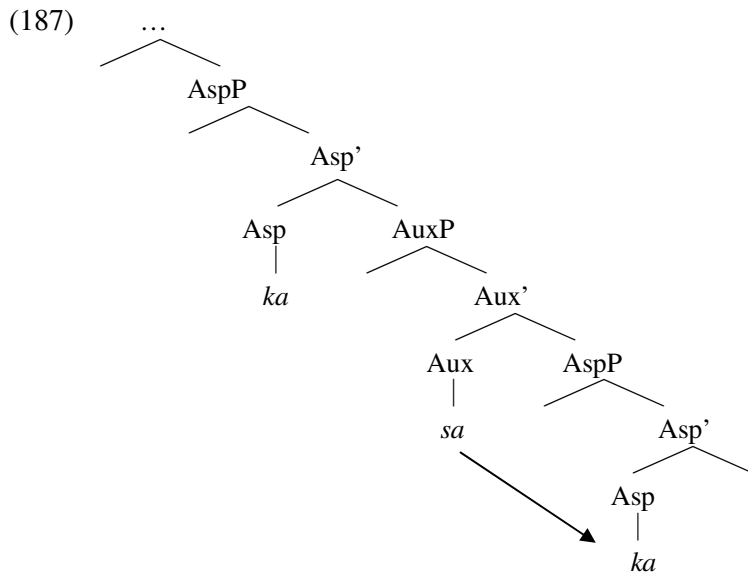
- (i) each item heads an independent AspP (cf. Cinque 1999, Aboh 2004);
- (ii) the items are merged in a single AspP by adjunction (Costa & Pratas 2004, Pratas 2004);
- (iii) The unit *ka ska* is stored as a single item in the lexicon and projects a single AspP in syntax.

I will immediately discard possibility (iii) because the several combinatories of tense-aspect markers, including the variation with *ka sa ka*, suggest that their individual components are all independently stored in the lexicon and acquire their (combined) semantic meaning in syntax. Scenario (iii) may be a scenario for the future, after full crystallization and a possible tendency towards a morphologically more complex language.

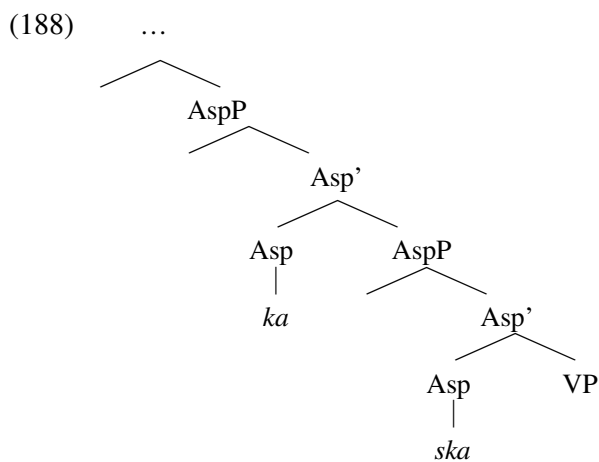
The strong version of the hypothesis in (ii), namely that TMA-markers and negation are recursively adjoined heads to T° in Capeverdean, fails to apply to Santome, one reason being the possibility of adverb stacking between T and Asp. However, under a weakened version of this proposal, one could assume that aspectual markers can adjoin as heads under AspP. This would look as follows:



Although I do not have any principled objection to this representation, I will adopt hypothesis (i) essentially for diachronic reasons, because it allows me to account for both *ka ska* and *ka sa ka*. Taking into account the diachronic development of the aspectual construction in question, I assume that representation (187) preceded (188):



In this configuration, as in (184) above, Aux merges with AspP and, as a consequence, the intermediate projection AuxP is eliminated from the structure, yielding the following representation:



I assume that the specifier of the lower AspP is opaque for adjunction or merge operations and that this tree also accounts for the *ka ka*-construction (cf. section 3.3.2.7), to which identical restrictions apply.

In sum, this section focused on the syntactic properties of the different aspect markers and I concluded that:

- aspect marker *ka* is a functional head (Asp°);
- auxiliary *sa* in the progressive construction is syntactically defective;
- most speakers interpret *ska* as a functional item that diachronically resulted from a post-syntactic process of phonological cliticization of an Aux° and an Asp°;
- the *sa ska*-construction corroborates the grammaticalization of *ska*;
- *ska* is presumably specializing for the progressive function of *ka*;

The next section focuses on the projection of tense but, owing to the intricate relation between tense and aspect, several aspects of this section will be restated.

3.3.6. Projecting tense

I will start this section with a brief consideration of the work of Comrie, whose insights on tense and aspect still prove to be very contemporary. This author dedicates a short sub-chapter to “aspect and time in tenseless languages”. His observations on isolating West-African languages carry over nicely to the tense-aspect system of Santome. Based on the case of West-African languages such as Yoruba, where the bare verb also conveys a perfective meaning and imperfectivity is expressed by preverbal marking, Comrie (1976: 82-3) states the following:

“In fact, in the absence of any contextual indication of time reference (e.g. a temporal adverbial), the Imperfective forms (...) are interpreted as referring to the present, while the Perfective forms (...) are interpreted as referring to the past (...). (...). Thus there is a close relationship between Imperfective Aspect and present time and between Perfective Aspect and past time, in these languages without tense markers.”

In agreement with Comrie, I assume that the concept of aspect in a language such as Santome is especially hard to dissociate from tense.⁷⁵ Similar observations have been made for the role of tense/aspect in language acquisition, where it has been emphasized that perfectivity is acquired before imperfectivity (e.g. Andersen & Shirai 1996)⁷⁶ and that past morphology does not pattern with the lexical semantics of aspect but rather

⁷⁵ Thus, earlier scholars working on the GGC, especially Valkhoff (1966) and Ferraz (1979), were partly correct when claiming that these languages are aspect-prominent languages.

⁷⁶ Within the Primacy of Aspect hypothesis, it is also claimed that perfectivity is first marked on punctual events (achievements) and then extends to accomplishment, activities and states.

with tense (Dietrich et al. 1995). Thus, considering the intimate link between creolization in general and L1 and L2 acquisition, and the well-attested transfer from West-African languages, the strong intertwining of temporal and aspectual information in the GGC is expected.

It is therefore tempting to establish a syntactic correlate between tense and perfectivity, on the one hand, and aspect and imperfectivity, on the other. Not surprisingly, authors concerned with mapping semantic information onto syntax in languages exhibiting Comrie’s “West-African” typology, i.e. where tense and aspect are separate categories from the verb, typically project TP and AspP (e.g. Aboh 2004, Durreleman 2000, Lefebvre 1998). Furthermore, following Michaelis (1993), I will argue that creole languages cannot do without tense in the core of their verbal system. In fact, it suffices to say that Bickerton’s anteriority or pluperfect marker, conveying exclusively temporal information, is found almost without exception in creole languages.

Comrie’s observations above readily extend to Santome, since a clear-cut aspectual opposition arises between the imperfective aspect marker *ka* (and in special cases also *ska*, *di* and \emptyset , as shown in the previous section), which is typically associated with present tense, and the so-called zero marker, which typically conveys past tense. Hence it follows that past marker *tava~ta* is actually the only core TMA-marker that is prominently tense-oriented. This section focuses on the syntactic properties of the latter marker in the following constructions:

(189) *Bô ta fada mu kuma fogon na bwa fa.*
 2SG TNS tell me that stove NEG good NEG
 ‘You had told me the stove wasn’t good.’

(190) *Jingu, non ta ka dêse ba poson.*
 Sunday 1PL TNS ASP go down go town
 ‘Sunday, we are going down to the city of S. Tomé.’

(191) *Ê na ta sêbê kuma kwa sa demono fa.*
 3SG NEG TNS know that thing is devil NEG
 ‘He didn’t know it was the devil.’

Example (189) represents the use of *tava~ta* as a pluperfect marker, (190) corresponds to the past imperfective and (191) illustrates the use of this marker with stative

predicates. The terms “tense marker” or “particle” are often used in creole studies and may be misleading in the sense that it seems as if it is *a priori* assumed that tense markers head (lexicalized) functional projections and, in this case, TP in particular. In what follows, I will therefore seek to answer the following two interrelated questions:

- (i) Do the instances of *tava~ta* in the constructions above correspond to a single item in the lexicon?
- (ii) How does *tava~ta* project in syntax?

To answer these questions, I will extensively discuss the properties of *tava* in the constructions above. Before discussing its syntactic properties, it should be noted that there is no evidence for a morpho-phonologically differentiated behaviour of *tava*. First, in any of the above constructions *tava* permits a reduced counterpart [*ta:*]. This reduction is not in any way constrained in either construction and should therefore be seen as mere variation. Second, there is diachronic evidence that supports a uniform evolution of *tava*. In spite of presenting a quite different description of TMA-markers compared to those described for contemporary Santome, Negreiros (1895: 327) shows that pluperfective *tava* already existed in 19th century Santome. I didn't find any examples of imperfective *tava ka* in his work, but only two instances of *stava* as a copula (see below) in a reproduction of poetry written by F. Stockler a decade or so earlier. In newspaper articles published in the early 1920s in *A Liberdade* and pamphlets written in the 1940s and 1950s, F. Bonfim on repeated occasions uses the construction *stava ka*, but in his pamphlets I was unable to detect any cases of *tava* as a pluperfect tense marker. Thus, not surprisingly, a few decades later, Ferraz (1979: 82-3) claims that both constructions present variation between [*tava*] and [*stava*], although I didn't attest the latter form in spoken contemporary language. It is also noteworthy that Ferraz does not mention the currently very productive short form *ta*.

In the previous section it was shown that aspect *ka* reduplicates. This morpho-phonological property is also available to the three types of *tava*-constructions, as illustrated in the following sentences:

- (192) *Mosu ta ta blôsê, lanta, ska ba golo migu dê.*
 boy TNS TNS angry get up ASP go look for friend POS
 The boy had gotten really angry, got up and went looking for his friend.’
- (193) *Zon tava tava ka kume pixi.*
 Zon TNS TNS ASP eat fish.
 ‘Zon was eating and eating fish.’ (kept eating)
- (194) *Zon tava tava sêbê kwa se.*
 Zon TNS TNS know thing SP
 ‘Zon really knew that.’

In sum, there is no diachronic or synchronic evidence that support two morpho-phonologically different forms of *tava*. Albeit inconclusive, these findings do not *a priori* rule out the possibility that there is a single item *tava* in the lexicon.

A number of independent syntactic tests can be applied to *tava* in the constructions under discussion. First, discontinuous negation has to embrace the full predicate, i.e. it cannot intervene between *tava* and the verb in either of the following *tava*-constructions.⁷⁷

- (195) *Zon *(na) tava *(na) ka kume pixi fa.*
 Zon NEG TNS NEG ASP eat fish NEG
 ‘Zon wasn’t eating finish.’
- (196) *Zon *(na) tava *(na) kume pixi fa.*
 ‘Zon hadn’t eaten fish.’
- (197) *Zon *(na) tava *(na) kônsê mina fa.*
 ‘Zon didn’t know the girl.’

Second, both constructions cannot be modified by temporally distinct adverbs.

- (198) **Onten, Zon tava ka ba poson oze.*
 yesterday Zon TNS ASP go town today
 (‘Yesterday, Zon was going to town today.’)

⁷⁷ In aspectual constructions in European Portuguese, for instance, negation can somewhat marginally intervene between the aspectual verb and the matrix verb (cf. Gonçalves 1996).

- (199) **Onten, Zon tava ba poson oze.*
 ('Yesterday, Zon had gone to town today.')

Considering that negation is dependent on tense, the two tests above show that both *tava*-constructions exhibit properties of monoclausal domains and therefore I assume that only a single instance of TP is projected in syntax (cf. Gonçalves 1999).

Third, the *tava*-constructions are able to embed a weather verb in (200) and (201), which is a property shared with, for instance, epistemic modals (202), but not with root modals (203).

- (200) *Ê ta sôbê.*
 'It had rained.'
 (201) *Ê ta ka sôbê.*
 'It was raining.'
 (202) *Ê toka sôbê.*
 'It should rain.'
 (203) **Ê mêsê sôbê.*
 3SG want rain

This difference can be explained simply by assuming that examples (200-202) are raising constructions, whereas (203) is a control structure. Since Ross (1969), it has been standardly assumed that root modals have a thematic role and therefore fall into the class of control verbs, whereas epistemic modals do not have a thematic role and thus pertain to the class of raising verbs. I will assume that *tava* has raising properties and is therefore defective from a thematic point of view. Note also that in the light of the weather verb test, essentially all the aspectual markers and aspectual verbs, as well as epistemic modals, can be considered raising verbs or perhaps to have gone through a diachronic stage where they were raising verbs.⁷⁸

Fourth, *tava* in either construction licenses VP-ellipsis, although the preferred strategy in these environments is stripping, as in the case of *sa* in the previous section.

⁷⁸ If one assumes, as sometimes is done, that the most deeply grammaticalized aspect marker *ka* derives from Portuguese *ficar* 'stay, remain', this would lend an additional argument in support of the raising hypothesis.

- (204) *Zon tava [kume pixi]_i, maji Maya na tava [-]_i fa.*
 Zon TNS eat fish but Maya NEG TNS NEG
 ‘Zon had spoken to Zwana, but Maya hadn’t.’
- (205) *Maya tava [ka kume pixi]_i, maji Zon na tava [-]_i fa.*
 ‘Maya was eating fish, but Zon wasn’t.’
- (206) *Zon tava [kônsê mina]_i, maji Zose na tava [-]_i fa.*
 ‘Zon knew the child but Zose didn’t.’

Similar results obtain for question-answer pairs.

- (207) Q: *Zon tava mata plôkô se?*
 Zon TNS kill pig SP
 ‘Had Zon killed the pig?’
 A: *Efan, ê tava.*
 ‘Yes, he had.’
- (208) Q: *Zon tava ka mata plôkô?*
 ‘Was Zon killing pigs?’
 A: *Efan, ê tava.*
 ‘Yes, he was.’
- (209) Q: *Zon tava kônsê mina?*
 ‘Did Zon know the child.’
 A: *Efan, ê tava.*
 ‘Yes, he did.’

It follows that the grammaticality judgments of my informants are more conclusive with respect to this construction than with respect to its past counterpart, the *sa ka*-construction of the previous section, for which the grammaticality judgments revealed some informant hesitation.

So far, the morpho-phonological and syntactic evidence suggests that the account of *tava* should be uniform. Basically, *tava* is a non-bound form that behaves like a raising verb involved in restructuring. However, there are also a number of facts that challenge a homogenous treatment of *tava* and warrant a split between past imperfective *tava ka*, on the one hand, and pluperfect *tava* and *tava* preceding stative

predicates, on the other. In essence, it can be shown that the former construction is less restructured than the latter two.

The first piece of evidence comes from pseudo-reflexive constructions (cf. section 2.6.3). As follows, I test whether pseudo-reflexives are able to attach to the immediate right of *tava* in both constructions and conclude that this is indeed possible in the *tava ka*-construction, in (210), but not in the pluperfect construction (211).⁷⁹

- (210) a. *N ta mu ka tason.*⁸⁰
 1SG TNS PSR ASP sit down
 ‘I was sitting down.’
 b. *N ta ka tason mu.*
 ‘I was sitting down.’
- (211) a. *N ta tason mu.*
 1SG TNS sit down PSR
 ‘I had sit down PSR
 ‘I had sat down.’
 b. **N ta mu tason.*

In (210a-b), the pseudo-reflexive is allowed to the right of *tava* and to the right of the verb *tason*, whereas only the a. sentence of (211) is grammatical.

Adverb placement is another good testing ground for these constructions and supports the proposed split. Adverbs such as *ten* ‘also, as well, too’, *leve-leve* ‘slowly’ or *pikina-pikina* ‘little by little’ are each able to intervene in the *tava ka*-construction, despite slight resistance from some speakers, but not in the other two constructions:

- (212) *?Zon tava {ten/leve-leve/pikina-pikina} ka kume pixi.*
 Zon TNS also/slowly/little by little ASP eat fish
 ‘Zon was also/slowly/little by little eating fish’
- (213) *Zon tava {*ten/*leve-leve/*pikina-pikina} kume pixi.*
 Zon TNS also/slowly/little by little eat fish

⁷⁹ Note that this test does not apply to stative predicates.

⁸⁰ Note that this also works for *sa* ‘to be’.

(i) *Ami sa mu tasondu ka kansa vida mu.*
 1SG be me sit-PP ASP rest life POS
 ‘I’m sitting by myself resting from my life.’

(214) *Zon tava* {**ten/*leve-leve/*pikina-pikina*} *kônsê Maya*.

Zon TNS also/slowly/little by little know Maya

Superficially, it seems as if the first sentence is somehow grammatical because of the presence of aspect marker *ka*. In fact, one could claim that the imperfectivity associated with *ka* in (215) allows some adverbs to intervene because *leve-leve* and *pikina-pikina* are actually imperfective in nature and would be ruled out by punctual predicates or states. The stative example, for instance, is fine upon conversion into a process:

(215) *Zon tava* {*ten/leve-leve/pikina-pikina*} *ka kônsê Maya*.

Zon TNS also/slowly/little by little ASP know Maya

‘Zon was also/slowly/little by little getting to know Maya.’

The adverb *ten*, however, shows that there is more to this than just stativity. If *ten* were stacked in sentence-final position, (213) and (214) would be fine. Thus, adverb placement suggests a structural difference between (212), on the one hand, and (213-214), on the other. At this point, it is necessary to reintroduce aspectual adverb *kwaji* ‘almost, about to’, which is exceptional in the sense that it occurs in all three constructions above, but alters their original interpretation.

(216) *Zon tava kwaji ka kume pixi*.

Zon TNS almost ASP eat fish

‘Zon was almost eating fish.’

(217) *Zon tava kwaji kume pixi*.

a. ‘Zon had almost eaten fish.’

b. ‘Zon was almost eating fish.’

(218) *Zon tava kwaji kônsê Maya*.

‘Zon was getting to know Maya.’

*‘Zon almost knew Maya.’

The specificity of *kwaji* in the past imperfective construction consists of the optionality of *ka*, which was also a property of the *sa ka*-construction in the previous section, as illustrated by examples (160-161). As a consequence, *kwaji* in the pluperfect structure lacking *ka* triggers an ambiguous reading between pluperfect and past imperfective. It

also follows from (218) that a stative predicate occurring with *kwaji* cannot be interpreted as a stative property predicate. Therefore, only the eventive reading with imperfective *ka* implicit is available. The conversion of states into processes was already discussed in section 3.3.3 and is foreseen in aspectual networks (e.g. Moens, 1987). Following the findings in section 3.3.3, I assume that stative property predicates should be set apart from other predicate types, including habitual states, due to their lack of internal structure. Syntactically, this difference can be accounted for by the fact that these predicates do not project AspP. But before discussing the syntactic structure of the tense-aspect markers, I will summarize the phono-syntactic properties of the constructions in Table 12 below. I also include the present progressive (*sa ka~ska~xka*), in order to determine whether and how it differs from the past progressive. The column corresponding to *tava+VP* stands for both dynamic and stative verbs, since they were shown to form a homogeneous syntactic class.

Table 12. Phono-syntactic properties of TMA-constructions.

	tava ka + VP	sa ka + VP	tava + VP
reduction to <i>ta</i>	✓	n.a.	✓
NEG tava / sa	✓	✓	✓
tava / sa NEG	x	x	x
disjoint temporal values	x	x	x
intervening PSR	✓	✓	x
intervening <i>kwaji</i> ‘almost’	✓	✓	✓
intervening <i>ten</i> ‘also’	?/✓	?/??	x
intervening <i>tan</i> ‘only’	?/??	?/x	x
VP-ellipsis	✓	?	✓
Null VP in Q-A	✓	?	✓
Embed weather verb	✓	✓	✓

The shadowed areas of this table indicate where the syntactic properties of the constructions in question differ. Very clearly, the *tava+VP*-construction (for pluperfect and past states) in the last column shows a tight relation between the tense marker and the predicate. In the progressive constructions, the relation between *tava/sa* and *ka+VP* is freer, which follows from adverb stacking and the placement of pseudo-reflexives. Although *tava ka* and *sa ka* share essentially the same syntactic properties, adverb placement and elliptic structures show that there is a difference in degree. I assume that

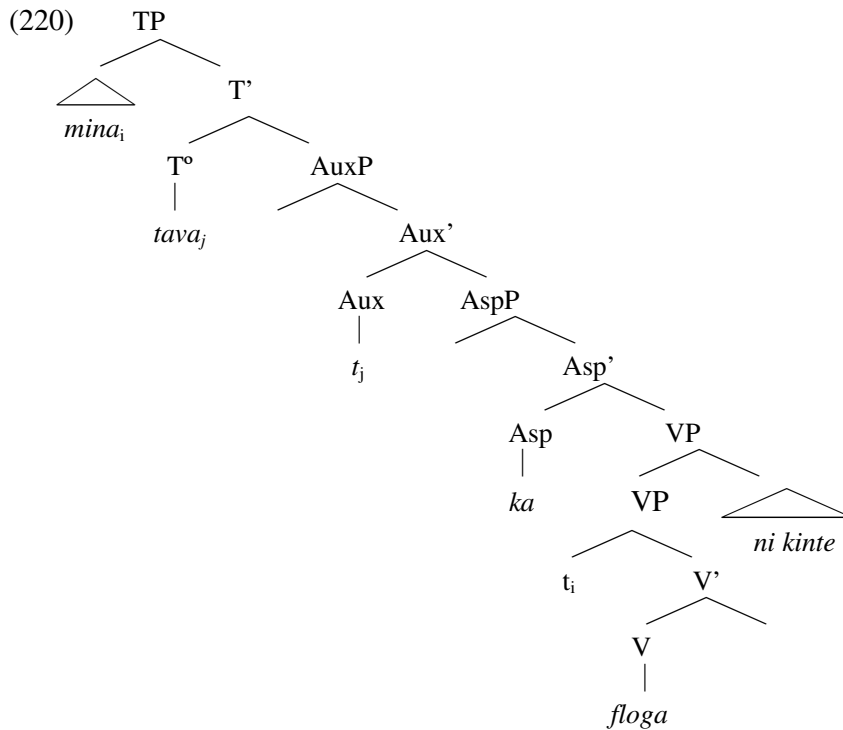
this is related to the fact that *tava* has retained more lexical features than *sa*. This is of course predictable if we consider that *sa ka* competes with deeper functionalized *ska~xka*. Despite the slight differences between the past and present imperfect, I will use the label Aux^o for both *tava* and *sa* in this construction. I assume that *tava* in the *tava*+VP-construction, on the other hand, heads TP. This distinction does not only follow the empirical findings, but it also has the advantage that it accounts for a basic split between, respectively, a temporal-aspectual category and a purely temporal category. Hence, the syntactic representation of the constructions under discussion needs to integrate the following findings:

- (i) All non-dependent predicates project TP in their structure;
- (ii) All non-dependent predicates project AspP, except for stative property predicates;
- (iii) The *tava*-constructions can be divided in two types with specific syntactic properties:
 - the *tava ka* VP-construction (past imperfective), with a looser relation between *tava* and the predicate,
 - the *tava + VP*-construction (pluperfect of dynamic verbs, past of stative verbs), with a tighter relation between *tava* and the predicate.

According to the findings above, I propose the syntactic representation in (220) for the *tava ka* VP-construction, as in example (219).

(219) *Mina tava ka floga ni kinte.*

‘The child was playing in the garden.’



In this raising structure, AspP is the complement of the AuxP headed by *tava*. It was shown that only a few adverbs that presumably adjoin to AspP can be stacked between both VPs. In the light of the data from VP ellipsis, I argue that *tava* raises to T°, where it enters a checking relation with the clausal subject, which undergoes long movement from [Spec, VP] to [Spec, TP].

This being settled, what determines that *tava* in this structure starts out as the head of a AuxP and not as T°? There are at least two considerations that would account for this derivation. First, it is important to remember that there is a rather obvious descriptive problem if one considers that in the progressive construction with *sa ka*, the present counterpart of the *tava ka*-construction, *sa* can be preceded by aspect marker *ka*, yielding the structure *ka sa ka*. As shown in section 3.3.2.6, this construction expresses a habitual or future progressive and requires an additional AspP layer on top of AuxP. In other words, accepting the assumption that *tava* in *tava ka* and *sa* in *sa ka* start out as the head of AuxP, *tava* is always able to raise to T, whereas *sa* can only raise to T when there is no intervening aspect marker (*ka*) to block head movement. This is independently confirmed by elliptic structures. Consider the following contrast:

(221) *Maya na ka sa ka tlaba ku ũa ngê fa; *Zon (ka) sa.*

Maya NEG ASP be ASP work with one person NEG Zon (ASP) be
'Maya is not going to work with someone; Zon is.'

(222) *Maya na {sa ka/ska} tlaba ku ũa ngê fa; ?Zon sa.*

'Maya isn't working with someone; Zon is.'

In (221), the contrastive clause licensing ellipsis is ungrammatical, irrespective of the presence of *ka*. Example (222), on the other hand, is considered grammatical with *sa ka* or *ska* when not preceded by *ka*.

The second argument in support of AuxP is related to examples such as the following:

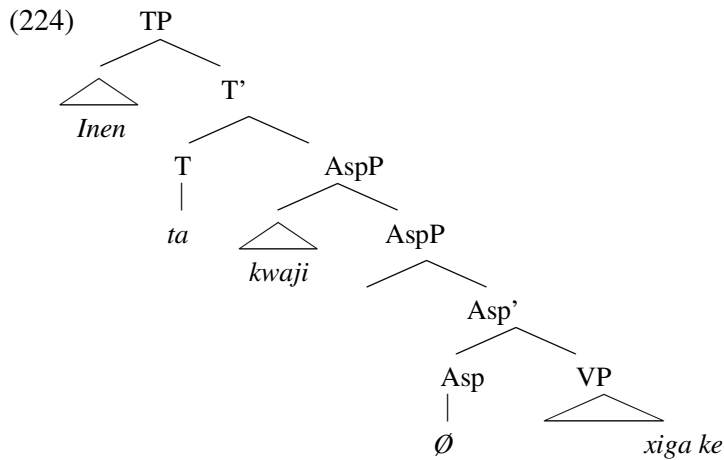
(223) *Inen ta kwaji xiga ke.*

3PL TNS almost arrive house

a. 'They had almost arrived home.'

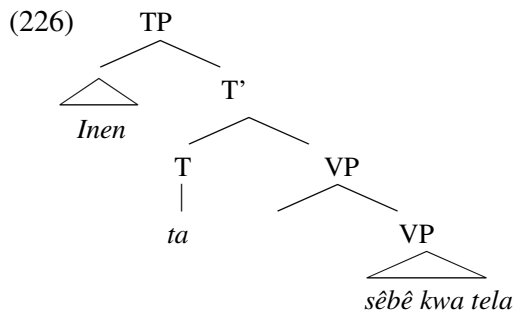
b. 'They were almost arriving home.'

Since I assume that *kwaji* is left-adjoined to AspP, I propose that the different interpretations have a separate counterpart in the syntactic structure that also takes into account the different syntactic properties of the past imperfective construction and the pluperfect construction. I therefore propose that the imperfective reading in (223b) corresponds to the tree in (220), with the sole difference that the head of AspP is null in this specific construction with *kwaji*. As for *tava* in the pluperfect construction, I assume it is merged as the head of TP. Reading (223a) thus corresponds to the structure illustrated in (224).



Finally, for the past of stative verbs I also propose that *tava* is merged in TP, since it was shown that it exhibits the same properties as pluperfect *tava* with dynamic verbs. The only structural difference that I propose concerns property predicates, on the one hand, and habitual states and stage-level predicates, on the other. More specifically, in section 3.3.3 I argued that property predicates do not project AspP, which is illustrated by the following example and the respective representation in (226).

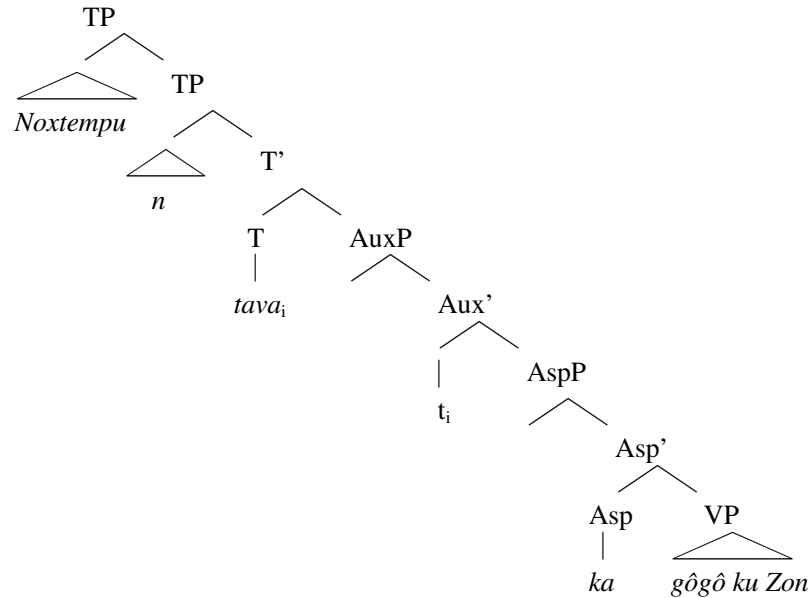
- (225) *Inen ta sêbê kwa tela.*
 3PL TNS know thing country
 ‘They knew the things about the country.’



Habitual states and stage-level predicates, on the other hand, project AspP, as shown in (227) and the respective tree in (228).

- (227) *Noxtempu n tava ka gôgô ku Zon.*
 formerly 1SG TNS ASP like with Zon
 ‘Formerly I liked Zon.’

(228)



The proposal outlined so far has the advantage that there is no need to store two or three different *tava* items in the lexicon. First, it was shown that *tava* is a lexical item that cannot be morpho-phonologically distinguished in any of the constructions in which it occurs. Second, the claim that *tava* in the past imperfective construction heads a VP, and in the pluperfect and stative constructions a TP, accounts for the syntactic differences found, such as adverb placement and pseudo-reflexivity. Hence, I assume that the lexicon contains a single lexical item *tava* with the feature [Past] which projects in the two proposed ways in syntax. Moreover, the projection AspP, or its absence with property states, is crucial in deriving the temporal interpretation of the sentence in which *tava* occurs.

Table 13. Typology of the *tava*-constructions.

<i>Construction type</i>	<i>Tense</i>	<i>Aspect</i>	<i>Outcome</i>
<i>tava ka</i> + V-construction	[Past]	[Imperfective]	past imperfective
<i>tava</i> + V _[-property state]	[Past]	[Perfective]	past before past
<i>tava</i> + V _[+property state]	[Past]	n.a.	past property state

The final temporal reading is determined by the syntax-semantics interface as a compositional feature of TP and AspP, if the latter is present. When *tava* heads a VP in the past imperfective construction, I assume that its [Past] feature becomes co-indexed with T by percolation. Therefore, there is no need for syntactic movement of this item.

This is similar to what Aboh (2004: 163) claims for Gungbe, with the difference that the temporal split is made between lexically realized future and non-future.

“When the feature [-future] is triggered, T° is realized by a null morpheme \emptyset that is controlled by an adverbial time specification or else by the default perfective aspect assigned to sentences involving no aspect marker.”

The assignment of null aspect to a dynamic verb in Santome mirrors the situation in Gungbe. Then, by merging *tava* as the head of TP, the combined feature of the perfective reading and the feature [Past] on *tava* trigger the pluperfect, i.e. Tense acts upon the perfective and the full structure now receives a past-before-past interpretation. A similar claim has been made for other languages as well, as follows from Durreleman’s (2000: 206) quotation with respect to Jamaican creole.

“Since [-stative] verbs have a default interpretation corresponding to the perfective reading, then the insertion of a [+past] tense marker *did* generally yields an anterior past interpretation.”

Note further that the use of *ka* as a future marker, in addition to its role as a marker of habitual and generic aspect, is not problematic under the current hypothesis. Besides the traditional problems related to the notion of future as a tense category, the idea of future is generally contextually determined (for instance by temporal clauses, adverbs, discourse, etc.) and is difficult to restrict to the aspect marker itself. Nevertheless, when *ka* receives a future interpretation, it never co-occurs with *tava* and therefore it assigns the temporal information to T. In this sense, Tense is therefore a compositional feature. Finally, the assumption that the only overt marker in T bears the feature [Past] underlines the fact that Santome, like most European languages, is best treated as a language that grammatically contrasts past, on the hand, and non-past.⁸¹

The proposed framework for aspect is also able to account for cases that apparently constitute a ‘mismatch’ between the semantics of the aspect marker and the temporal interpretation of the clauses in which they occur (cf. section 3.3.4). Consider the following two examples:

⁸¹ On this distinction, see Comrie (1985: 44).

- (229) *Ê d'e kwa kume, ê pê lôpa d'e, p'ê bixi. Mosu*
 3SG give-3SG thing eat 3SG put clothing give-3SG for-3SG dress boy
ska dispidji di mana.
 ASP say goodbye of sister
 'He gave him food, he provided clothes of his for him to wear. The boy was
 saying [lit. is saying] goodbye to his sister.'
- (230) *Soku alê ten da blandu kuma nen tlêtlêxi mina ska kaza ê.*
 Then king also give news that 3PL three-RED girl ASP marry EMPH
 'Then the king spread the news that all three girls were going to marry.'

In these examples, the progressive sentence receives past reference, in spite of the fact that imperfectivity in Santome is by default related to present tense. As mentioned in section 3.3.4, it follows that in these cases a non-local tense-chain is established whereby usually TP with linear precedence provides the value [+Past] to the following TP or TPs. This means that clauses in Santome possess two ways of anchoring temporal reference:

- (i) Locally, by a TP-AspP-chain;
- (ii) Non-locally, by a TP-TP chain.

Finally, the fact that aspect marker *sa ka~ska~xka* is commonly bound by external tense constitutes additional evidence for the fact that *sa* does not head TP.

In the previous section it was suggested that even the non-contracted present progressive construction (*sa ka*) shows signs of being a grammatical chunk that has visible effects on the grammaticality of adverb placement and ellipsis. In this section it was shown that the past imperfective *tava ka~ta ka*, the past counterpart of the *sa ka*-construction, is much less constrained with respect to identical tests. It is important to note that there is no direct parallel that can be drawn between *tava ka* and *sa ka* and *ta ka* and *ska*, because *ta ka* is not a contracted form and its properties are the same as those of *tava ka*. Nevertheless, grammaticalization would pass through a stage in which short *ta ka* becomes the dominant construction. Therefore, it is interesting to check whether the synchronic and diachronic distribution (cf. Tables 14 and 15) suggest that a change may be taking place. Table 14 shows the contemporary findings.

Table 14. Number of occurrences of the past imperfective in the spoken corpus.

	<i>number of occurrences</i>	<i>% (approx.)</i>
<i>tava ka VP</i>	36	51
<i>ta ka VP</i>	34	49
Total	70	100

These numbers suggest that there may well be no reason to believe that *tava ka* is currently giving place to *ta ka*. However, it is possible that a diachronic change is indeed occurring if we consider that in Bonfim's pamphlets⁸² the only structure used six times is *stava ka*. Ferraz (1979:82-3) claims that there is variation between *tava* and *štava*. The form *stava* occurs once in my corpus in a copula construction of an elderly speaker. In sum, the diachronic data possibly suggest that *stava~xtava* was at variation with *tava*, whereas the current variation exists between *tava* and *ta*. That the difference between *tava* and *ta* is not numerically relevant also follows from cases where these markers immediately precede the VP, which I have divided in states and dynamic verbs, as shown in the following table.

Table 15. Number of occurrences of the *tava*+VP-construction in the spoken corpus.

		<i>number of occurrences</i>	<i>% (approx.)</i>
<i>tava VP</i>	state	11	21
	dynamic	9	18
<i>ta VP</i>	state	23	45
	dynamic	8	16
Total		51	100

In sum, it can be concluded that *tava* and *ta* appear to be used in rather equal proportions synchronically but it also followed that a diachronic shift took place. Another salient fact is the relatively low number of occurrences of the different past constructions compared to the imperfective constructions. For instance, the total number of occurrences of *sa ka* and *ska~xka*, 1461, stands in shear contrast with the 70 instances of *tava ka~ta ka*, which represents less than 5%. A part of the explanation for this discrepancy comes from the fact that most of the spoken corpus consists of

⁸² I did not find instances of the past imperfective in Negreiros (1895).

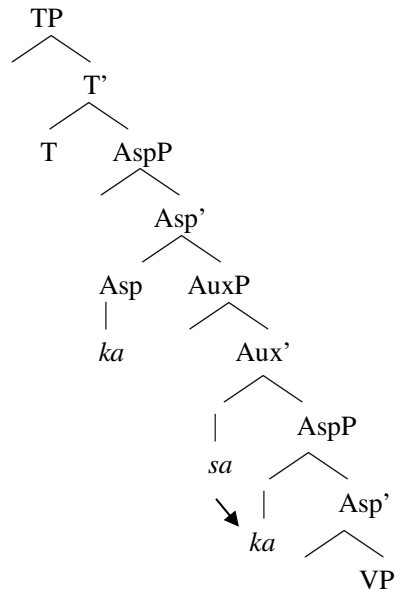
narratives that use present narrative tense and also from the fact that aspect is often anchored by tense outside the clause it occurs in. Moreover, temporal and conditional clauses, which are very frequent, typically lack overt tense (*tava*).

3.3.7. Summary

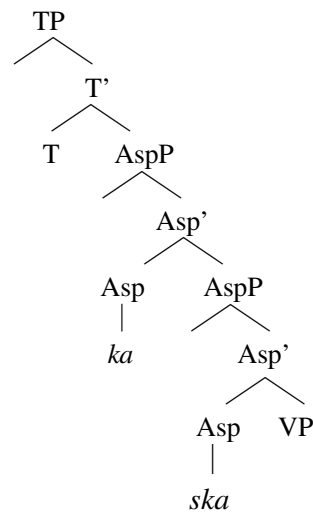
The sections on tense and aspect show that in Santome these two categories cannot easily be dissociated, given the especially tight relations between perfectivity and past, on the one hand, and imperfectivity and present, on the other. Following the previous literature, I consider *ka*, *sa ka~ska~xka*, *tava~ta* and the non-marked verb the core tense and aspect material. I have shown that *ka* is the most functional element of these markers and behaves as a bound morpheme. It was also observed that the progressive marker *sa ka~ska~xka* involves some complexity owing to the fact that we are dealing with a contracted form. Although generally considered synonyms by native speakers, the full form *sa ka* and the contraction *ska~xka* exhibit a number of syntactic differences. I assumed that the contracted form can be either analyzed as an instance of a post-syntactic merger, i.e. phonological cliticization, or as an instance of fusion, i.e. a fully grammaticalized form that is stored as such in the lexicon. The latter scenario, which accounts for such forms as *sa ska* and *ka ska*, would represent a more advanced diachronic stage of the language. The main feature of *tava~ta*, I argued, is [Past]. The syntactic behavior of this lexical item shows that it still preserves lexical features, in contrast with *sa ka~ska* and especially *ka*, which are almost depleted, or are fully depleted, of lexical features.

To account for the empirical findings, I have proposed three different syntactic labels, namely AuxP, TP and AspP. In addition to *ka* and *ska*, the zero marker for perfectivity and items such as *di* (as in *sa di*, for instance) head AspP. In specific cases (e.g. *sa ska*, *ka ka*), I claimed that a double aspectual layer best accounts for the data. I further argued that only *tava~ta* in pluperfective constructions and with stative predicates can head TP. I further assumed that *tava* and *sa* in the present and past imperfective constructions, respectively, head AuxP. The distinction between TP and AuxP corresponds to the distinction between a temporal projection and a temporal-aspectual projection respectively. The following tree structures summarize the syntactic typology of tense and aspect marking in Santome:

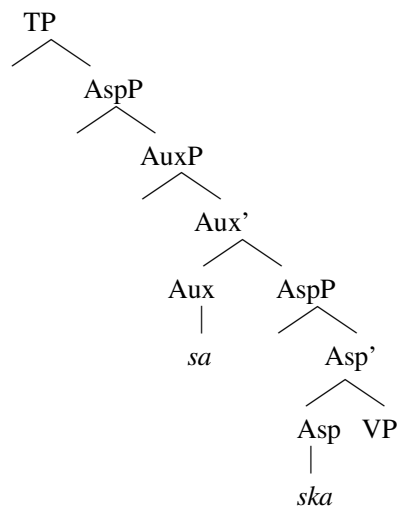
(231) Habitual/future progressive



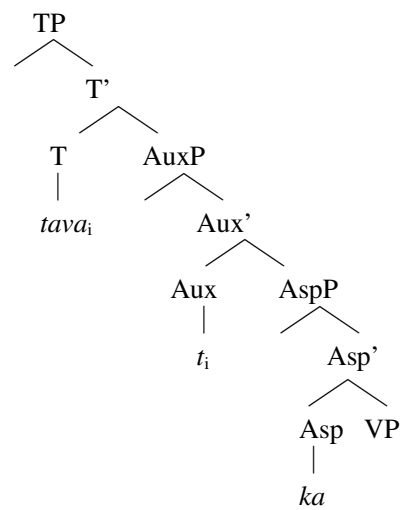
(232) Habitual/future progressive

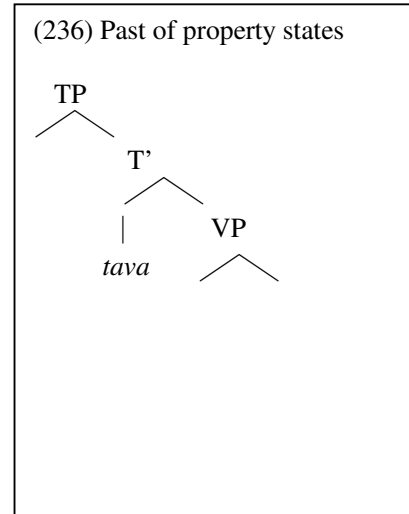
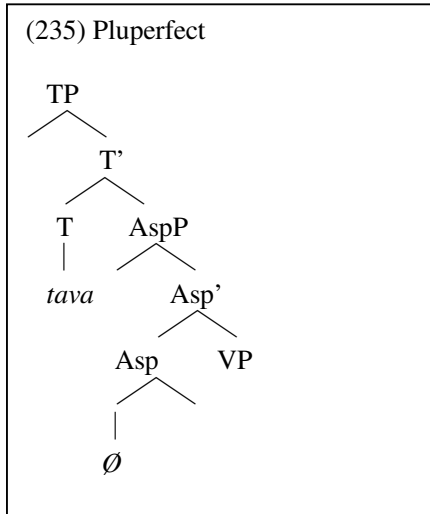


(233) present (perfect) progressive



(234) Past imperfective





The trees in (231) and (232) represent the full and the contracted structures of the habitual/future progressive. I assume that for some speakers *sa* is still an AUX that can merge post-syntactically with *ka*. However, in most cases, I believe, *ska* is already analyzed as a single grammatical item heading an AspP, as per (232). In a present (perfect) progressive, in (233), *sa* is an AUX and *ska* sits in Asp°. In the past imperfective construction, represented in (234), with *tava ka*, *tava* is generated in AuxP but can be raised to TP. Although the present counterpart in (231) bears significant similarities to this structure, the occurrence of the first *ka* in *ka sa ka* blocks raising of *sa* to T°. The pluperfect in (235) operates with a Ø aspect marker with *tava* directly generated in T°. The same applies to the past interpretation of property states, in (236), but this structure differs from all the other structures by lacking an AspP. Since *tava* always carries the feature [Past], the aspectual information, when present, sets the stage for the final temporal interpretation of the clause. An AspP headed by an imperfective marker (*ka*, *ska*) typically triggers a present reading. *Tava*, with scope over the AspP, then contributes the feature [Past]. When AspP is headed by a zero marker, as in (235), a default perfective (past) reading is obtained. The presence of *tava* then contributes [Past], yielding, as predicted, a past-before-past.

3.3.8. Mood and modality

The considerable attention paid to tense and aspect by scholars working on Santome and GGC in general stands in sheer contrast with the discussion on mood in these languages, despite the central role of this category in Bickerton's (1975, 1981) work on TMA-systems. I believe that the main reason behind this silence with respect to mood derives from the simple fact that no specific mood marker has been identified in Santome's core TMA-system. In section 3.3.8.1, I will argue that such a mood marker does however exist and projects in the TMA-domain. Section 3.3.8.2, in turn, discusses evidence in support of specific modal items in the low left periphery and a correspondent syntactic projection

3.3.8.1. Mood in the TMA domain

This section concerns the properties of the mood marker *ká*. This item differs from aspect marker *ka* by the fact that it carries high tone, which is easily perceptible for native speakers. To distinguish between both markers, I have chosen to mark high tone with an acute accent. Although not mentioned in the literature on Santome, Maurer (1997) refers to a similar *ká* in closely-related Lung'ie and labels it a counterfactual marker. Although this marker occurs most frequently in counterfactual environments, namely conditional clauses, this designation does not cover some of the uses below. Therefore, I propose the label "mood marker". In (237), (238) and (239), *ká* occurs respectively in a purpose clause, a clause embedded under a predicate of doubt and a conditional clause.

(237) ... *pa* *bô* *ká* *be* *ku* *ami*.

for 2SG MOOD go with 1SG

'(...) so you could go with me.'

(238) *Sun* *diskunfya* *ya* *ê* *ká* *pô* *sa* *jingantxi* *ku* *môlê*.

He suspect that 3SG MOOD can be giant REL die

'He suspected that it could be the giant who died.'

- (239) *Xi non d'ola se na ká tava ka da ku ngê*
 if 1PL of-hour SP NEG MOOD TNS ASP give with people
tamen fa, mo ngê d'oze ka vivê?
 adult NEG how people of-today ASP live
 'If we back then wouldn't have gotten along with the adults, how would today's
 people live?'

Mood marker *ká* exhibits all the properties of a functional morpheme and occupies a fixed position between the subject and the VP. Example (239) is especially interesting, because it shows the precise locus of this marker in the preverbal functional complex, namely in between negation (*na*) and tense (*tava*). Note further that, unlike the case of aspect marker *ka*, the consonant [k] of *ka* does not obligatorily adopt the point of articulation of weak 1sg *n* (cf. section 2.6.1), as illustrated here:

- (240) a. *Xi n ká/gá be....*
 If 1SG MOOD go
 'If I would go...'
 b. *N ga/*ka ba ke*
 I ASP go house
 'I go home.'

Since this position of the mood marker does not comply with Bickerton's claim that creole preverbal systems typically exhibit the linear order T-M-A, Maurer (1997) proposes the designation "MTA-system" in order to account for the high position of the mood marker. MTA is therefore a distinctive feature of the GGC.⁸³ In the light of the functional behavior of *ká* and its fixed position in clause structure, I assume that this marker heads a MoodP(hrase) within the extended VP.

It should also be noted that the \emptyset -marker often conveys mood. This follows from examples such as the following, which are identical to contexts where *ká* was shown to appear, such as purpose clauses and conditional clauses.

⁸³ I do not have any information on the existence of a high mood marker in Fa d'Ambô.

- (241) *Bi da mu ãa kopu d'awa pa n Ø da sun se.*
 come give me a cup of-water for 1SG give man SP
 'Bring me a cup of water to give to that man.'
- (242) *...punda xi bô Ø bêbê pasa, bô ka fika bêbêdadu xinku ja.*
 because if 2SG drink surpass 2SG ASP stay drunk five day
 '...because if you drink too much, you will be drunk for five days.'

It follows that this Ø-marking is unrelated to the Ø-marker for perfectivity, since it refers to a hypothetical event situated in the future. I therefore assume that MoodP as proposed above may also be headed by this Ø-marker, in the same way that TP and AspP can house Ø-markers, i.e. items with a phonetically empty output. I further assume that MoodP only projects when the structure requires it. In these cases, if MoodP is the topmost projection in the clause, the clausal subject moves from [Spec,VP] to [Spec,MoodP].

3.3.8.2. High modality

The discussion above focused on mood within the classic TMA system. However, in addition to the mood markers and modal verbs, the latter of which I will not discuss here, Santome also exhibits modal items in pre-subject position. The properties of these markers, *sela* 'must' and *minhon* 'better', suggest that Santome possesses a low left-peripheral projection that encodes the speaker's attitude towards the proposition.

3.3.8.2.1. *Sela*

The first item I will discuss, *sela* (*pa*) 'must', was already studied by Ferraz (1979: 84-5), who labeled it a "particle of obligation". Ferraz was particularly interested in describing the uses of *sela*, its semantics and its etymology, but not so much in its syntax. As shown in (243-244), this necessity marker occurs at the front of the clause.

- (243) *Sela non fla santome ben fladu.*
 Must 1PL speak Santome well spoken
 'We must speak good Santome.'
- (244) *Sela pa ê pê kalu fôgô.*
 must for 2SG put kalu fire
 'You must put the *kalu* (traditional stew) on the fire.'

Differently from the *pa* that introduces complement and adjunct clauses (e.g. purpose clauses), the status of *pa* in the above type of clauses is optional and doesn't block tense marking. These facts follow from the contrast between (245) and (246):

- (245) *Sela (pa) n tava fe kwa se.*
 must for 1SG TNS do thing SP
 'I should have done that.'
- (246) *N tava mēsê *(pa) bô (*tava) fe kwa se.*
 1SG TNS want for 2SG (TNS) do thing SP
 'I wanted you to do this.'

In (245) tense is anchored by the matrix clause and cannot occur in the embedded clause. *Sela* in (246), on the other hand, does not exhibit any clausal properties. It is a fully frozen item in the sense that it can neither take (expletive) subjects nor preverbal marking, as shown in (247). Moreover, it cannot stand alone (for instance as an answer), as in (248), and the clause selected by these modal particles cannot be moved across it, as in (249), thus contrasting with complement and adjunct clauses. The absence of these properties is illustrated in the following examples.

- (247) *(*Ê) (*tava) (*ka) sela bô pê kalu fôgô.*
 3SG TNS ASP must 2SG put kalu fire
- (248) Q. *Sela ê pê kalu fôgô?*
 must 3SG put kalu fire
 'Must he put *kalu* on the fire.'
- A. *Efan, (*sela).*
 'Yes, he must.'
- (249) **[(Pa) bô fla] sela.*
 (for) 2SG speak must

It is also important to note that this modal do not behave like adjuncts. This follows not only from the lack of any discourse break but also from its fixed position in the clause, which follows from its position with respect to appositives and adverb placement and the subject (*non*) in examples (250-251). Example (252) further shows that the presence of *pa* still yields an ungrammatical sentence.

(250) a. *Onten, ê fla, non ba ple.*
 yesterday 3SG say 1PL go beach
 'Yesterday, he said, we went to the beach.'

b. **Sela, ê fla, non ba ple.*

(251) a. *(Oze) Zon ba omali (oze).*
 Today Zon go sea today
 '(Today) Zon went to the beach (today).'

b. *(Oze) sela (*oze) Zon ba omali (oze).*
 today must today Zon go beach today
 '(Today) Zon must go to the beach (today).'

(252) a. **Sela pa, ê fla, non ba ple.*

b. **Sela, ê fla, pa non ba ple.*

Focus marker *so* has to precede *sela*, as shown in (253). Focusing to a position below it is precluded, in (254).

(253) *Kinte_i so sela pa bô bali [-]_i.*
 garden FOC must for 2SG sweep
 'You must sweep the GARDEN.'

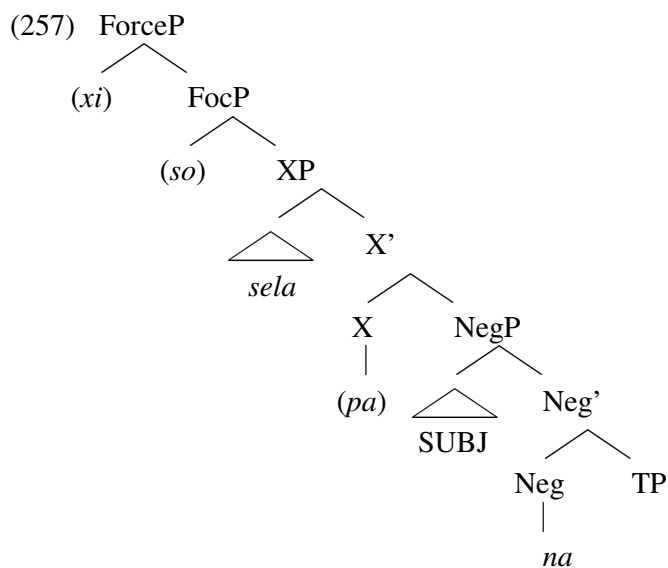
(254) **Sela pa kinte so bô bali.*
 must for garden FOC 2SG sweep

In Chapter 2 it was shown that strong 1sg *ami* is fully grammatical when it precedes negation marker *na*. This also applies when *ami* does not occur in sentence-initial position, as illustrated in the conditional clause in (255). When *ami* follows *sela*, however, the sentence becomes ungrammatical. Only *n* can occur in this case, as follow from (256).

(255) *Xi n/ami na fla fa...*
 if 1SG NEG speak NEG
 'If I don't speak, ...'

(256) *Sela n/*ami na fla fa.*
 must 1SG NEG speak NEG
 'I must not speak.'

If *ami* were a DP in the standard subject position, such as [Spec,TP] or [Spec,NegP] in the sentences above, it should be grammatical. The fact that *sela* bans *ami* from the subject position in (256) supports the hypothesis that *sela* is located in a low left-peripheral functional projection, whereas *ami* fills the specifier of a projection in between a high conjunction such as *xi* ‘if’, and the projection hosting *sela*. Assuming Rizzian cartography, for instance, *xi* would head ForceP. The strong pronoun is arguably the specifier of FocP, especially if one considers that negation, as in the sentences above, may impose focusing effects on subjects. The abbreviated structure in (257) spells out this possibility.



Note that *xi* and *so* are of course the heads of ForceP and FocP, respectively. I suggest that *sela* and *pa* establish a Spec-Head relation in XP. As for the nature of XP, note that several authors, for instance Cinque (1999) or Speas (2004), have argued for a refined syntactic treatment of modal features such as evidentiality or evaluation, which can be grammatical categories cross-linguistically. I propose that the XP in Santome is a Modal Phrase (ModP) and in the next section I will provide evidence that *sela pa* is not the only modal element that sits in this projection.

3.3.8.2.2. *Milhon*

In addition to *sela*, *milhon* ‘(it is) better’ in (258-260) also conveys deontic modality but exhibits a lower degree of commitment than *sela*. Like *sela*, *milhon* may occur with *pa*, and also with *xi* ‘if’:

- (258) *Milhon n ga môlê mu, punda tudu kwa se ka kaba dê.*
 Better 1SG ASP die PSR because all thing SP ASP end of-3SG
 ‘I had better die, because everything comes to an end.’
- (259) *Milhon pa bô mala mu anzu pê tlaxi.*
 better for 2SG tie 1SG baby put back
 ‘You had better tie me the baby on my back.’
- (260) *Milhon xi n bê wê dê.*
 better if 1SG see eye POS
 ‘I had better see his eyes.’

Structurally, it can be readily shown that *milhon* occurs in the same position as *sela*, as will be demonstrated below. This follows for instance from adverb placement, focus and TMA-marking in respectively (261), (262) and (263).

- (261) (*Oze*) *milhon* (**oze*) *bô bêbê vin se.*
 (today) better (today) 2SG drink wine SP
 ‘Today you had better drink this wine.’
- (262) [*Jaka*]_i *so milhon bô subli* [-]_i.
 Jacktree FOC better 2SG climb
 It is the jacktree you had better climb.’
- (263) *Milhon (pa) n gá tava môlê.*
 Better (for) 1SG MOOD TNS die
 ‘I had better die.’

Like *sela*, *milhon* is a crystallized form that cannot take a subject, negation, and TMA-marking, nor be moved in syntax. Note also that *sela* and *milhon* cannot co-occur, in (264), which is expected in the light of the different modal reading associated with each item.

- (264) a. **Sela (pa) milhon (pa) bô ba ke.*
 must (for) better (for) 2SG go house
- b. **Milhon (pa) sela (pa) bô ba ke.*

Note finally that this syntactic pattern also carries over to likely recent borrowings⁸⁴, as illustrated by the deontic item *importanti* in the following example:

- (265) *Importanti a kume bêbê.*
 Important IMP eat drink
 ‘It is important that one eats and drinks.’

3.3.8.3. Summary

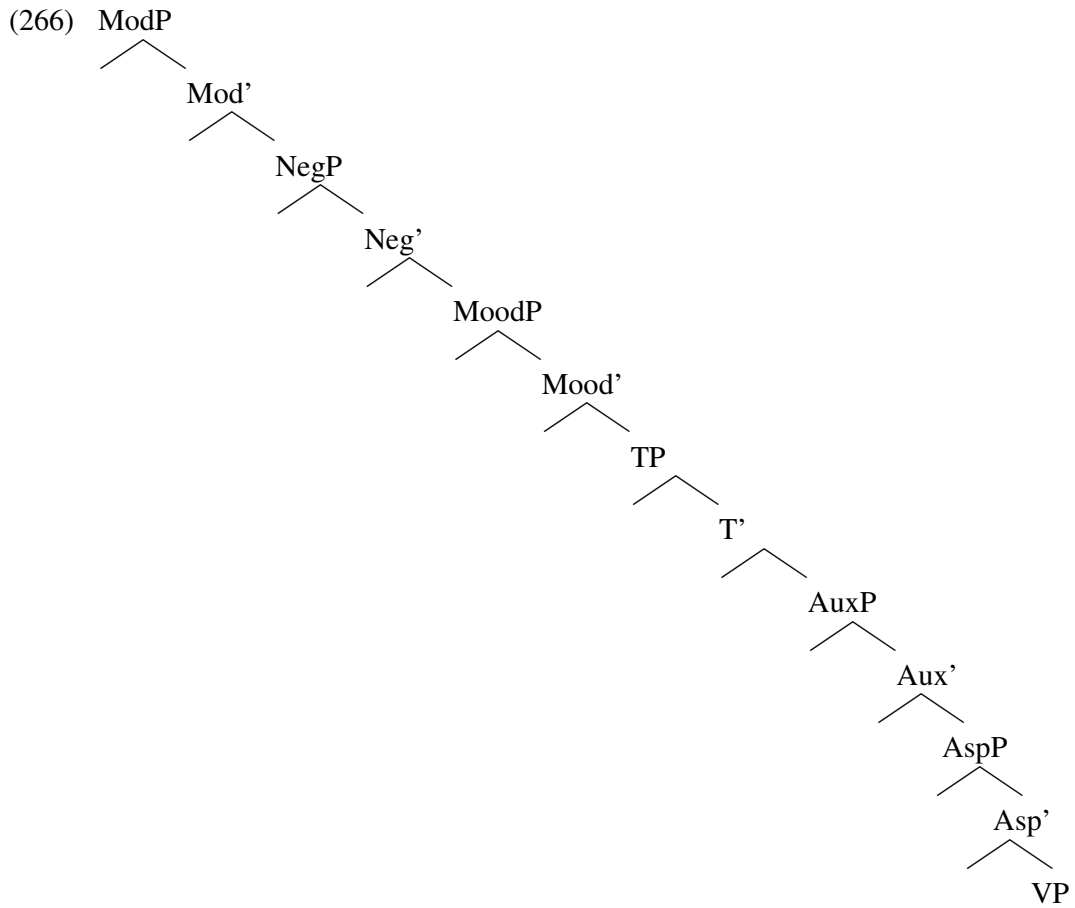
I have shown that Santome exhibits a functional projection headed by mood marker *ká*, which should not be mistaken with aspectual *ka*, given its specific phono-syntactic properties. Since the mood marker precedes the tense marker *tava~ta*, I assume that Santome has an MTA-system, following Maurer’s (1997) proposal for Lung’ie. In addition to describing the mood marker, I provided evidence for another functional projection hosting at least several deontic modal markers.⁸⁵ In the light of the data, I assume that this projection is the lowest projection in the Santome left-periphery. The discussion of the core TMA-markers in this chapter has resulted in the following functional structure:

⁸⁴ See Ferraz (1979:23-4) on recent borrowings and their non-conformity with the phonological rules of the language.

⁸⁵ In Chapter 1, section 9.2.4, it was shown that *palêsê* is modal item. Note, however, that it has at least two properties that prevent it from occurring in ModP, namely the fact that it can occur above a FocP and adverb placement:

- (i) *Palêsê oze ê ska dwentxi.*
 seem today 3SG ASP be ill
 ‘It seems that he is ill today.’
 (ii) *Palêsê pixi so ê ska kume.*
 seem fish FOC 3SG ASP eat
 ‘It seems that is is fish he’s eating.’

Comparing these examples with the teste for *sela*, for instance, it follows immediately that *palêsê* sits higher in the clause.



I have claimed that TP always projects in the preverbal domain. The same applies to AspP, except when property predicates are mapped onto syntax. I further assume that the remaining functional projections only project in response to empirical motivation. This also may be said of the second AspP-layer, which I did not represent in the structure above. Note that I have already included the preverbal negation marker in the tree above. In the next chapter, I will provide substantial evidence for NegP in this position and I will refine the above structure in the light of the findings concerning the syntax of the clause-final negation marker *fa*.

4. NEGATION IN SANTOME

4.1. Introduction

Default clausal negation in Santome involves two negation markers, *na* and *fa*. The former, to which I assign the label Neg1, occurs in preverbal position whereas the latter, which receives the label Neg2, occurs typically in clause/sentence final position.

Cross-linguistically, the negation pattern with two markers is highly marked (e.g. Kahrel 1996) and has received different labels, such as double negation, discontinuous negation, split negation or bipartite negation. In addition to these labels, I will also use the label double-headed negation, since it will be argued that Neg1 and Neg2 are syntactic heads. An important feature that underlies these labels is that in a language with discontinuous negation like Santome negative items co-occur with both negation markers without canceling negation.

This chapter develops as follows. Section 4.2 contains a description of the patterns of discontinuous negation in Santome. Section 4.3 discusses the status of both negation markers. Next, in section 4.4, I will review the analyses that have been proposed for the specific negation patterns in other languages. Finally, in section 4.5 I will propose a new analysis building especially on data from adjunct placement and coordination.

In the Appendix to this chapter, I will focus on negative concord (NC), polarity items and diachronic and comparative aspects of negation in the GGC.

4.2. Sentence negation: the data

Although Santome is known in the literature for its discontinuous negation patterns, which have often been referred in connection with negation patterns in other creole or vernacular varieties such as Palenquero or Vernacular Brazilian Portuguese, it is actually the case that this language exhibits three distinct patterns for sentence negation:

- (i) the default (or unmarked) pattern Neg1...Neg2;
- (ii) a marked pattern with Neg1 only;
- (iii) a marked pattern with Neg2 only.

The following sections will provide a descriptive account of the syntax and semantics of each of these patterns.

4.2.1. Contexts with Neg1...Neg2

As mentioned, default negation in Santome requires two negation markers, *na* and *fa*, whose individual properties will be discussed in more detail in sections 4.3.1 and 4.3.2. Neg₁ occurs before the TMA-markers (cf. Ch. 3) whereas Neg₂ surfaces in a strongly clause-final position.

4.2.1.1. Simplex sentences

In normal declarative sentences, Neg₁ follows the subjects and precedes TMA-material or auxiliary verbs and Neg₂ occurs in sentence-final position, as illustrated in the following examples:

- (1) *Men dê na tava sêbê fa?*
Mother POS NEG TNS know NEG
'Didn't his mother know?'
- (2) *Ê na sê piska fa.*
3SG NEG know fish NEG
'He can't fish.'

In the case of a few modal auxiliaries, such as *pô* 'can, may' and *toka* 'should', Neg₁ may follow the verb in order to provide a specific modal reading. Note that in these cases *na* can also precede the modal.

- (3) *Ê pô na sêbê lê fa, ê tê valôr muntu.*
3SG can NEG know read NEG 3SG has value much
'Even though she may not be able to read, she has a lot of value.'
- (4) *Ê toka na fla fa.*
3SG should NEG speak NEG
'He should not speak.'

Imperative clauses typically lack an overt subject and therefore Neg₁ occurs in sentence-initial position.

- (5) *Na fe mu vlegonha fô!*⁸⁶
 NEG do 1SG shame NEG-EMPH
 ‘Don’t make me ashamed.’

Neg2, which behaves as a bound morpheme with respect to the material to its immediate left, occurs to the right of the verb and its complements. This is exemplified by a subject relative, a relativized direct object and a double object construction with a relativized Theme, in (6-8) respectively.

- (6) *Mina mosu ku na ka ngosta di ta ke fa ska lêlê*
 small boy REL NEG ASP like to be house NEG ASP accompany
san ni tlaxi.
 lady in back
 ‘A boy who doesn’t like to stay home accompanies his mother.’
- (7) *Ê na ka bila konsê xitu ku kwa sa nê fa.*
 3SG NEG ASP turn know place that thing be in-3SG NEG
 ‘He doesn’t recognize the place where the thing is.’
- (8) *Ê na fada mana dê kwa ku kod’e fa.*
 3SG NEG tell sister POS thing REL wake-3SG NEG
 ‘He didn’t tell his sister about what woke him up.’

PPs and adverbials that follow the verb typically occur to the left of Neg2 as well, even postverbal sentence-level adverbs such as *amanhan* ‘tomorrow’.⁸⁷

- (9) *Nê ãa ngê nê ladron na ka poto ala ku ope fa.*
 not one person not-even thief NEG ASP step there with foot NEG
 ‘Nobody, not even a thief, enters that place.’
- (10) *Bô na ka kume muntu fa.*
 2SG NEG ASP eat much NEG
 ‘You don’t eat much.’

⁸⁶ *Fa* and emphasis marker *ô* normally contract into *fô*.

⁸⁷ When sentence-level adverbs occur in preverbal position, they obligatorily precede Neg1.

- (11) *Punda kwa ku ê ka fla oze, ê na ka pô bi*
 Because thing KU 3SG ASP say today 3SG NEG ASP can come
fl'e amanhan fa.
 say-3SG tomorrow NEG
 'Because what he says today he may not say tomorrow.'

Note that in all the cases above the complements and adjuncts occur *obligatorily* to the left of Neg2, i.e. Neg2 is clause-final. There are, however, constituents that behave more freely with respect to Neg2. First, this is the case of certain temporal adverbials, which may occur to the right of *fa*.

- (12) *A na da mu kume fa jina plaman.*
 IMP NEG give 1SG food NEG since morning
 'They haven't given me food since the morning.'
- (13) *N naxi laba boka fa n naxi kume fa antê minda d'ola se.*
 1SG NEG wash mouth NEG 1SG NEG eat NEG until measure of-timeSP
 'I haven't brushed my teeth yet nor have I eaten until that moment.'
- (14) *Soku n mêsê p'ô fada mu kwa fe vin se na da fa sêxi*
 So 1SG want for-2SG tell 1SG thing do wine SP NEG give NEG six
mêji za.
 month already
 'So I want you to tell me why the palm tree isn't giving wine for six months already.'

The specificity of temporal constructions with durative 'until', 'since' and 'for' with respect to event structure has been observed in the literature (e.g. Kamp & Reyle 1993: 628-34). Note, however, that in its spatial use, adjuncts headed by *jina* and *antê* have to occur to the left of *fa*.

- (15) *A na ka be antê gala dê fa.*
 IMP NEG ASP go until heart POS NEG
 'They don't go until the heart of it [the palm tree].'

- (16) *A na fe bixa jina ple fa.*
 IMP NEG make cue since beach NEG
 ‘They didn’t line up from the beach.’

This is expected, since the spatial use does not interfere with the event structure of the clause. I will discuss the syntax of temporal *jina* and *antê* in section 4.5.1.1.3.

Sentence-level particles expressing insistence/emphasis occur obligatorily to the right of Neg2, as illustrated in the following examples:

- (17) *Sun na tôlô fa ô!*
 He Neg1 silly Neg2 EMPH
 ‘He (formal) is not silly!’

- (18) *N na sa klupadu fa ê!*
 1SG Neg1 be guilty Neg2 EMPH
 ‘I’m not guilty!’

- (19) *Kyê avo, punda Dêsu avo, na da mu fa fan!*
 EXCL grandma because God grandma Neg1 give 1SG Neg2 EMPH
 ‘Oh, please grandma, don’t beat me!’

The discussion of the relation between the final negation marker and emphatic elements, especially *fan*, can be found in the Appendix to this chapter, section 4.2. Finally, vocatives, although usually in sentence-initial position, are also found in final position, following Neg2.

- (20) *Kwa na sa dôtôlô fa, papa mu.*
 Thing Neg1 be doctor Neg2 friend POS
 ‘That is not a doctor, my friend.’

In sum, *fa* is strongly clause-final in simplex sentences. Only sentence-level particles, vocatives and a very restricted range of temporal adjuncts occur to its right.

4.2.1.2. Complex sentences

This section describes the behavior of Neg2 with respect to different types of clausal domains. Unlike simplex sentences where *fa* occurred almost exclusively to the right of all the material, it will be shown that the structural position of Neg2 in complex sentences is dependent on the type of clause-linking, and determines scope relations.

4.2.1.2.1. Embedding with Neg2 in sentence-final position

When *na* occurs in a matrix clause selecting a clausal complement, *fa* occurs invariably at the end of this complement clause, i.e. in sentence-final position. The following examples list the main types of complement clauses in Santome. Complementizer selection is determined by the main predicate. Declarative and epistemic verbs typically occur with complementizer *kuma*⁸⁸, verbs of volition and inquiry typically select *pa*, verbs of doubt *xi*, and factives, generally *mo* or *punda mo*. This can be observed in respectively (21) to (24). Note further that in some cases declaratives and epistemics allow for null complementizers, as illustrated in (25).

- (21) *Ome se na fla kuma ê sa kunhadu bô fa.*
man SP NEG say that 3SG be brother-in-law POS NEG
'That man didn't say he's your brother-in-law.'
- (22) *Sun na mêsê pa sun ba nala ku mosu sun se fô.*
3SG NEG want for 3SG go in-there with boy POS SP NEG-EMPH
'He (formal) doesn't want to go there with his son.'
- (23) *Maji n na sêbê xi n ga nganha ala fa.*
but 1SG NEG know if 1SG ASP arrive there NEG
'But I don't know if I get there.'
- (24) *Ê na poda mu mo ku n da mina dê ku kwa fa.*
3SG NEG forgive 1SG way REL 1SG give child POS with thing NEG
'He didn't forgive me for beating his child.'
- (25) *N na kunda bô sa Dêsu fa.*
1SG NEG think 2SG be God NEG
'I dind't think you were God.'

⁸⁸ Note that *kuma* can be reduced to *ku*, for instance when material is extracted out of the complement clause.

In all the examples above, the matrix predicate is negated, whereas the subordinate predicate isn't. This raises of course the question as to what clausal domain Neg2 belongs. From example (26) it readily follows that syntactically speaking *fa* does not belong to the embedded clause, since this item is stranded when the embedded clause is fronted.

- (26) [*Kuma ê sa kunhadu bô*]_i, *ome se na fla* [-]_i *fa*.
 that 3SG be brother-in-law POS man SP NEG say NEG
 'That he is your brother-in-law, the man didn't say.'

Simultaneously negating the main and the embedded clause results in a double occurrence of Neg1 and a single instance of Neg2, as shown in (27a). Neg2 cannot occur twice in final position. But fronting of the embedded clause in (27b) shows that both domains are, in fact, independently negated:

- (27) a. *Ome se na fla (*fa) kuma ê na sa kunhadu bô fa (*fa)*
 man SP NEG say (NEG) that 3SG NEG be brother-in-law POS NEG
 'The man didn't say he isn't your brother-in-law.'
 b. [*Kuma ê na sa kunhadu bô fa*]_i, *ome na fla* [-]_i *fa*.
 'That he isn't your brother-in-law, the man didn't say.'

The absence a double Neg2 in final position can presumably be assigned to a rule of haplology. However, note that, according to Post (1997), Neg2 in closely related Fa d'Ambô can be doubled and is even compulsory with complement clauses:

- (28) *E na bi fa e na ske bi=f=uf.* (Fa d'Ambô; Post 1997: 300)
 3Sg NEG ANT say 3SG NEG IRR come=Neg2=Neg2
 'He did not say he would not come.'

When the complement clause is negated, the discontinuous pattern is obligatory as well.

(29) *San fla: ti, sa kinte se ku san fada mu pa n na ba*
 She say friend be garden SP REL she tell 1SG for 1SG NEG go
floga nê fa.

play in-3SG NEG

‘She said: my friend, it’s the garden that I [she] told you not to go play in.’

(30) *N fada san kuma n na sêbê fa.*
 1SG tell 3SG that 1SG NEG know NEG

‘I told her I didn’t know.’

In section 4.2.2. it will be shown that, unlike negated complement clauses, negated adjunct clauses introduced by complementizer *pa* do not always realize Neg2.

In addition to the complement clauses above, several other constructions allow this type of ‘long distance’ placement of Neg2. This happens in the following constructions.

Serial verb constructions

(31) a. *N na tê jêlu di ligi tlega nansê fa.*
 1SG NEG have money to lift up hand over 2PL NEG
 ‘I don’t have money to give to you.’

b. *Nansê na ka subli ba ôbô fa.*
 2PL NEG ASP go-up go forest NEG
 ‘You don’t go up to the forest.’

Temporal final clauses

(32) *Zon na kume plumê zo pa bêtê fa.*
 Zon NEG eat first then for drink NEG
 ‘Zon didn’t drink before eating.’

Circumstantial negative clauses

(33) *Ê na ka nda sê pa ê da topi fa.*
 3SG NEG ASP walk without for 3SG give trip NEG
 ‘He doesn’t walk without tripping.’

Comparative and conformative clauses

- (34) *Zon na sa maxi lôngô dôkê manu dê fa.*
3SG NEG be more tall than brother POS NEG
'He isn't taller than I am.'
- (35) *Tempu sa kentxi, vin na mêsê to mo ê ka to fa.*
weather be hot wine NEG want drip way 3SG ASP drip NEG
'The weather is hot, the palm wine doesn't want to drip the way it usually does.'

*Final relative clauses*⁸⁹

- (36) *Firminu soku na da mu plastiku pa n dêsê ku ê fa.*
Firminu FOC NEG give 1SG plastic-bag for 1SG descend with 3SG NEG
'Firmino didn't give me a plastic bag to go down (to town) with.'

Causal clauses

- (37) *Punda xi n tê kabelu blanku, na sa punda n sa fitisêlu fa.*
because if 1SG have hair white NEG be because 1SG be sorcerer NEG
'Because if I have white hair, it doesn't mean I'm a sorcerer.'

In all these examples, only the final position is available for *fa*. Although I will not exhaustively discuss the syntactic structure of all these examples, the general observation is that the embedded clause must occupy a relatively low position in the structure.

As for serial verb constructions, in (31) above, I argued in Hagemijer (2000) that in Santome VP₂ in SVCs can be analyzed as adjuncts to a lower AspP.⁹⁰ This correctly predicts that both VPs that form a complex predicate are in the scope of a single NegP. In fact, one of the widely accepted diagnostics for SVCs is the impossibility of independently negating any of the VPs that constitute these constructions.

Examples (32) and (33) are presumably instances of adjunction of a CP to VP, (34) an adjunction of a CP to AdjP and (36) adjoins a final relative (*pa n dêsê ku ê*) to

⁸⁹ Another designation for these clauses is purpose relative clause because of the combination of a relative and a purpose clause (e.g. Palmer 1986: 178).

⁹⁰ Note that VP₂ in serializing languages typically does not carry tense. Sometimes VP₂ may carry aspect (e.g. Santome and Saramaccan), but in other cases, e.g. Haitian, all the verb modifying material is concentrated on VP₁.

the Theme argument of the double object verb *da* ‘give’. Finally, the causal clause in example (37) is a VP-adjunct. In section 4.5 it will be argued that causal adjuncts may adjoin to different syntactic projections.

Therefore it follows that the clausal adjuncts illustrated above attach relatively low in the structure. The highest domain for adjunction in the examples is AspP. None of the loci of adjunction constitute a barrier for the strictly final placement of Neg2. The data covered so far lead to the following preliminary generalization:

- (38) In Santome, selected material and adjuncts to AspP or lower projections are within the scope of Neg2.

4.2.1.2.2. Embedding with Neg2 in clause-final position

So far it followed that *fa* is strongly sentence-final. In this section it will be shown that there is also a wide array of contexts where it cannot surface outside the clause that houses Neg1. The following structures either block Neg2 or are independent domains for negation:

Negative coordination and enumeration

- (39) *Inen na ka fla fa nê inen na ka pô fl'e fa.* (Bonfim)
 3PL NEG ASP speak NEG nor 3SG NEG ASP can speak-3SGNEG

‘They don’t speak nor are they allowed they speak.’

- (40) *Bô na da mu niku kume fa, bô na da mu niku bêbê*
 2SG NEG give 1SG nothing eat NEG 2SG NEG give 1SG nothing drink
fa.

NEG

‘You don’t give me anything to eat or to drink.’

- (41) *Kaso se na tê opê fa, na tê mon fa, na tê*
 dog SP NEG have leg NEG, NEG have forefoot NEG NEG have
dentxi fa, maji ê ka môdê pasa.
 tooth NEG but 3SG ASP bite surpass

‘That dog doesn’t have backfeet, forefeet and teeth, but it has a mean bite.’

It follows from the examples that negative coordination can either be syndetic or assyndetic. The fact that the subject is repeated shows that we are dealing with a high

adjunction structure, above Neg1, which explains why each conjunct has to be negated independently. The type of enumeration (or multiple coordination) in (41) is presumably an instance of adjunction at subject level, whereby the overt subject (*kasô*) is able to c-command the following empty subject positions. Since negative coordination is an important test-case with respect to the placement of *fa*, I will return to these structures in section 4.5.1.3.

Adversative clauses

- (42) *Mali na ka kaba ku mali fa, sela ben.* (Daio, 2002: 22)
 bad NEG ASP finish with bad NEG, only good
 ‘Bad doesn’t put an end to bad, only good does.’
- (43) *Amôlê seku na pega fa, so monhadu!* (Bonfim)
 love dry NEG stick NEG only wet
 ‘Dry love didn’t stick, only wet love.’

The *sela/so*-constructions obtaining contrastive readings are arguably elliptic in nature and should therefore occupy a higher adjunction site, crucially above negation. The elliptic structures are illustrated in (44-45).

- (44) *Mali na ka kaba ku mali fa*, [_{CP} *sela* [_{IP} *ben* [_{VP} ~~*ka kaba*~~]]].
- (45) *Amôlê_i seku na pega fa*, [_{CP} *so* [_{IP} *t_i monhadu* [_{VP} ~~*ka pega*~~]]]

Causal clauses

- (46) *Mina na ka pô kaza ku pobli fa, punda pobli sa pobli.*
 Girl NEG ASP can marry with poor NEG because poor be poor
 ‘The girl cannot marry a poor guy because a poor guy is a poor guy.’
- (47) *Maji non na ka kula môlê fa, plukê myole Sun Govenadô ska*
 but 1PL NEG ASP cure death NEG because now Mr. Governor ASP
*sama non ku amôlê, ku sosegu, pa non bi nganha djêlu.*⁹¹
 call 1PL with love, with calmness, for 1PL comeearn money
 ‘But we don’t cure the deaths, because now the Governor is calling us with love,
 calmness, to come and earn money.’ (Bonfim)

⁹¹ *Plukê* is arguably an older form that occurs frequently in Bonfim’s texts. I didn’t find it in my own data and it is not mentioned in the literature on Santome.

Reason clauses

- (48) *Kuma fala d'alê na tê lefeta fa, jingantxi toma*
because speech-of-King NEG have importance NEG Giant take
plinxeza mêtê pê saku (...)
princess put put bag
'Because the word of the king has no importance, the giant put the princess in a bag.'

Conditional clauses

- (49) *Xi ê na bê faka fa, ê na bêbê vin fa.*
if 3SG NEG see knife NEG 3SG NEG drink wine NEG
'If he doesn't find the knife, he won't drink wine.'

Concessive clauses

- (50) *Dedu di ngê pô na bwa fa, a na ka kot'e zuga*
finger of person can NEG good NEG IMP NEG ASP cut-3SG throw
buta fa.
throw NEG
'Even if somebody's finger is not good, you don't cut it off and throw it away.'
(Daio 2002: 56)

Temporal-spatial clauses

- (51) *Ola ben na sen fa, baleladu ka stivi.*
when good NEG exist NEG so so ASP suffice
'When good doesn't exist, so so is enough.'
- (52) *Sun Faxiku na lega vesu fa antê ê bi da tudu kwa se*
Mr. Faxiku NEG abandon verse NEG until 3SG come give all thing SP
di 1953.
of 1953
'Mr. Faxiku didn't stop writing until all those things of 1953 occurred.'

- (53) *Ami, jina n sa ai, n na t êndê fala di nê ùa ngê fa,*
 1SG since 1SG be here 1SG NEG hear speech of not one person NEG
jina a zuga mu pê glêntu saku.
 since IMP throw me put inside bag
 ‘Me, since I’m here, I haven’t heard nobody’s speech, since I was put in the bag.’

Final clauses

- (54) *Soku manda sun fla pa tudu ngê di tudu flêgêja ku naxi bi tlabá*
 So therefore 3SG say for all people of all village REL NEG come work
fa, pa tudaxi dêsê ni fin di mêji se plama, senfata.
 NEG for all come down at end of month SP morning without delay
 ‘So, therefore he says that everybody from all the villages who hasn’t come to do the work yet, for all of them to come down at the end of the month in the morning, without delay.’ (Bonfim)

Temporal adverbials

- (55) *Maji ê na tê trêxi dja fa, sode ten mal’e plôvya zôgô.*
 but 3SG NEG have three day NEG soldier also tie-3SG because game
 ‘But less than three days ago, the soldiers also tied him up because of the game.’
- (56) *Na tada fa, pe dê lentla.*
 NEG last NEG father POS enter
 ‘Shortly after, his father entered.’

4.2.2. Contexts without Neg2

Santome presents a number of constructions in which *fa* does not occur alongside Neg1. These constructions have in common a lack of commitment to the truth condition of the negation.

When negation occurs inside *pa*-clauses that express purpose, Neg2 does generally not occur.

- (57) *N fuji fala pa fala na lêlê mu.*
 1SG flee rumor for rumor Neg1 follow me
 ‘I fled from the rumors, so they wouldn’t follow me.’

- (58) *Milhon pa bô na b'êlê.*
 better for 2SG NEG see-3SG
 'You had better not see him.'
- (59) *Kwidadu pa manjinga se ku n sa ku ê na subli mu.*
 careful for aggressive person SP REL 1SG be with 3SG NEG go up 1SG
 'Be careful that the aggressive person in me doesn't take over.'
- (60) *Palêsê pa a na fla kum'ê tê klupa, ê kaboka dê.*
 Seem for MP NEG speak that-3SG have guilt 3SG keep silent POS
 'Apparently for them not to say he's guilty, he kept silent.'

Example (57) is a standard purpose clause, whereas (58-60) are purpose clauses embedded under a modal expression (admonitions, doubt). This syntactic constraint sensitive to the modal interpretation of these clauses has an effect similar to the use of subjunctive morphology in identical constructions in Romance languages. Note that Neg2 *fa* can occur in these constructions, but according to my informants it is construction-sensitive. The *kwidadu*-construction, for instance, is more generally accepted without Neg2 than the *milhon*-construction.

Despite the tendency for standard purpose clauses such as (57) to occur without Neg2, I found some exceptional cases where Neg2 occurred. In (61) below, the occurrence of Neg2 is arguably related to the fact that the contrastive clause introduced by *maji* 'but' has a canceling effect on the hypothetical status of the purposive clause.

- (61) *Mwala ska golo p'ê fuji pa n na b'êlê fa, maji n*
 Woman ASP search for-3SG flee for 1SG NEG see-3SG NEG but 1SG
b'êlê.
 see-3SG
 'The woman was trying to flee so I wouldn't see her, but I did see her.'

Indirect and direct interrogative constructions without Neg2 are also common. In direct interrogatives, the question marker *an* often shows up in the final position.

- (62) *Kloson na ka dwê sun an?*
 heart NEG ASP hurt 2SG INT
 'Doesn't your heart ach?'

However, Neg2 is not excluded in interrogatives and its presence or absence yields a contrast in the presuppositions held by the speaker:

- (63) a. *Bô na bê nadaxi di bisu ni kabêsa mu fa?*
 b. *Bô na bê nadaxi di bisu ni kabêsa mu an?*
 2SG Neg1 see nothing of animal on head POS {NEG/INT}
 ‘Haven’t you seen any animals on my head?’

In yes/no questions such as (63), the presence of *fa* represents the speaker’s presupposition of a negative answer based upon his knowledge of the world⁹², whereas the b. sentence with question marker *an*, which could optionally be omitted as well, is neutral with respect to the speaker’s presupposition. The following sequence reinforces the idea that *fa* in questions carries a negative expectation and insists upon prior discourse.

- (64) A: *N fla an: bô kônsê mu?*
 1SG say EMPH 2SG know 1SG?
 ‘I said: do you know me?’
 B: *Não.*
 no
 A: *Bô na kônsê mu fa?*
 2SG Neg1 know 1SG Neg2
 ‘Don’t you know me?’

When a verbal complement functions as an indirect interrogative clause, Neg2 may or not occur, as illustrated in the two following pairs:

- (65) a. *Inen na sêbê ola Zon mata plôkô.*
 3PL NEG know when ZOn kill pig
 ‘They don’t know when Zon killed the pig.’
 b. *Inen na sêbê ola Zon mata plôkô fa.*
 ‘Idem.’

⁹² Note that in this case *fa* triggers an effect similar to negative tags.

- (66) a. *N na sêbê mo ê nganha ke.*
 1SG NEG know how 3SG arrive house
 ‘I don’t know how he got home.’
- b. *N na sêbê mo ê nganha ke fa.*
 ‘Idem.’

According to my informants, the context with *fa* provides a higher degree of certainty with respect to what is being negated. Like the interrogatives and the *pa*-constructions above, *fa*’s presence or absence is related to doubt or uncertainty and therefore belongs to the domain of epistemic modality.

Furthermore, exclamatives are also commonly produced without Neg2. Example (67) represents the speaker’s desire, (68) is a case of irony and (69-70) are instances of expletive negation, where negation is not actually semantically contentful (e.g. Portner & Zanuttini, 2000).

- (67) *Ê na pô nganha posta se!*
 3SG Neg1 can win bet SP
 ‘No way could he win that bet!’
- (68) *Jina n sen, n naxi b’êlê!*
 since 1SG exist 1SG Neg1 see-3SG
 ‘I have never seen him in my whole life!’
- (69) *Inen na sêbê mo inen nganha ke!?*
 3PL Neg1 know how 3PL arrive house
 ‘They know how they got home!’
- (70) *Kê kwa! N na fada nansê kwa se!*
 what thing 1SG Neg1 tell 2PL thing SP
 ‘What!?! Didn’t I tell you so!’

Finally, in written productions I was able to find a few proverbs where conditional clauses lack Neg2:

- (71) *Tudu kasô sôsô, xi ê na môlê ni flankotxi, bobom ka*
 all dog stray if 3SG NEG die by rifle ? ASP
ving'e sawôji.
 take revenge-3SG health
 'Every stray dog, if it doesn't die by rifle, ?the plague will get him.'
 (Daio 2002: 83)

Due to their hypothetical nature, conditional clauses fit the typology of constructions without Neg2. Although it was shown in 2.1.2.2 that conditional clauses exhibit the full discontinuous pattern, proverbs often represent older stages of the language, which can be considered an indication that the exclusive Neg1 pattern used to be more widespread in subjunctive environments.

Kahrel (1996) provides evidence for the non-factual status of additional negation markers cross-linguistically. His finding rests upon the fact that these markers, generally corresponding to Neg2 in Santome, often have multiple functions that altogether can be described as irrealis. It was shown, however, that modern Santome does not fit in to this typology, because Neg2 in this language is a fully grammatical negation marker and discontinuous negation is the unmarked pattern. Rather than its presence, it is the absence of *fa* that triggers marked interpretations. The next section underlines that Santome exhibits a continuum of negation patterns that ranges from these arguably more conservative patterns to the most innovative patterns, which consist of Neg2 exclusively.

4.2.3. Contexts without Neg1

Some instances of contrastive negation lack the preverbal marker (Neg1), but it should be noted that this happens only sporadically. This type of negation, which occurs either with *fa* but frequently with its emphatic counterpart *fô*, requires an appropriate discourse trigger, as follows from the following examples:

- (72) *Ni glêntu d'awa? Sabi kyê nê fô!*
 in inside of-water key fall in-3SG NEG-EMPH
 'In the water? The key didn't fall in there!'

- (73) *Ola ku ngê ka saka nda kansadu, ê toka bêbê awa fô.*
 whenKU people ASP ASP walk tired 3SG must drink water NEG-EMPH
 ‘When people are tired, they shouldn’t drink water.’

This pattern hinges upon the denial of immediately prior discourse and constitutes yet another piece of evidence for the link between Neg2 and intensification. Note in this respect that, in both the above examples, the emphatic form is employed. Another piece of evidence for the marked reading of the exclusive final marker comes from the use of Neg2 to negate constituents contrastively.

- (74) *Karu fa!*
 Car Neg2
 ‘Not the car!’
- (75) *Zon ka fla, glita fa!*
 Zon ASP speak shoutNeg2
 ‘Zon speaks; he doesn’t shout!’

4.2.4. Summary

The data above show that the placement of *fa* is sensitive to the syntactic environment and intimately linked to the scope of negation. It was shown that the type of embedding is crucial to determine the structural position of *fa*. It followed that *fa* belongs syntactically to the same clause as preverbal negation marker *na*.

The three main patterns described above suggest that Santome exhibits the stages that are typical of Jespersen’s cycle: exclusive preverbal negation represents the most conservative pattern, whereas exclusive final negation is presumably the most progressive pattern. In between these marked patterns, discontinuous negation stands for the unmarked negation. However, some care is required, since comparing the contemporary data with older sources and the other GGC does not suggest that a shift towards an exclusive final pattern is taking place. In section 4.4 of the Appendix to this Chapter, more details on the diachrony of negation in Santome and the GGC can be found.

4.3. The status of the negation markers in Santome

This section will investigate the morphosyntactic status of *na* and *fa*, in particular whether these elements are heads or specifiers.

4.3.1. The status of Neg1

Following Zeijlstra's (2004: 141) assumption that preverbal negation markers are cross-linguistically heads with syntactic (free elements or particles) or morphological status (affixes or clitic-like elements), it is expected that Neg1 in Santome is a head.

Before entering the details of this issue, it should be noted that Santome exhibits a tripartite system of preverbal negation markers. In addition to the standard preverbal negation marker *na*, there are two other negation markers that occur in preverbal position as well, namely *naxi* 'not yet' and *nanta(n)*⁹³ 'no longer, not anymore, never (again)', as illustrated in the following examples:

(76) *Ê na tê salu fa.*
3SG NEG have salt NEG
'It doesn't have salt.'

(77) *Bô naxi tava nansê ten fa.*
2SG NEG TNS born also NEG
'You weren't born yet either.'

(78) *Oze so n nanta(n) ka dansa fa.*
Today FOC 1SG NEG ASP dance NEG
'Today I won't dance anymore.'

Although it is not clear whether *tan* and *xi* are historically derived from X^os or XPs, I assume that these forms are fully lexicalized. It can be readily shown that these three negation markers compete for the exact same syntactic position and share all their properties. The arguments in support of this claim are the following:

⁹³ Post (1997, 2000) always analyzes *tan* in closely related Fa d' Ambô as an iterative marker (IT) in her glosses, as in the following example:

(i) *Fosyi na tan sa iai f.* (Fa d' Ambô, Post 1997: 312)
strength NEG IT be here NEG
'The strength is not there anymore.'

Given the significant similarities between the negation patterns in both creoles, the examples she provides strongly suggest that *tan* forms a complex negation marker with *na*, as in closely related Santome.

(i) they occur lower than the subject and higher than the verb or its TMA markers

- (79) *N*(na/naxi/nanta) tava (*na/naxi/nanta) ka (*na/naxi/nanta) fla fa.*
I (NEG) TNS (NEG) ASP (NEG) speak NEG
'I wasn't speaking.'

(ii) They are mutually exclusive

- (80) a. **Zon na nantan fla fa.*
b. **Zon na naxi fla fa.*
c. **Zon nantan na fla fa.*
d. **Zon nantan naxi fla fa.*

(iii) They typically occur together with Neg2 (*fa*) in default environments

- (81) *N {na/naxi/nanta} fla *(fa).*
'I {didn't speak / haven't spoken yet / didn't speak anymore}.'

(iv) They all license N-words

- (82) *Sun na ka pô fe nadaxi fa.*
3SG NEG ASP can do nothing NEG
'He couldn't do anything.'
- (83) *Maji n naxi tê nadaxi fa.*
But 1SG NEG have nothing NEG
'But I don't have anything yet.'
- (84) *Mosu nantan fla ku nê ũa ngê fa.*
Boy NEG speak with not one person NEG
'The boy didn't speak with anybody anymore.'

(iv) *They cannot occur in isolation (e.g. as answers to questions)*⁹⁴

(85) Q: *Zon ka be poson?*

Zon ASP go town

Does Zon go to the city of S. Tomé?

A: **Na/*naxi/*nantan.*

(vi) *They cannot occur in negative stripping.*

(86) *Zon be, maji (*na/*naxi/*nantan) Maya fa.*

Zon go but NEG Maya NEG

‘Zon went, but not Maya.’

In sum, it is evident that these three markers exhibit the same syntactic properties and compete for the same structural position in Santome’s sentence architecture. Several of the proposed tests relying on the Head Movement Constraint (Travis, 1984) in order to determine the status of negation markers, such as (long) clitic climbing or blocking of verb movement (Zanuttini 2001) fail to apply because, as I argued, Santome lacks syntactic clitics (cf. Ch. 2) and verb movement (cf. Ch. 3).

A note is in place with respect to the fact that, despite the presence of two negation markers for sentence negation, the three preverbal negation markers have negative force by themselves. French *ne*, for instance, does not carry negative force by itself.

(87) *...pa a pô na sêbê kuma ê mata bisu.*

for IMP can NEG know how 3SG kill animal

‘...so they wouldn’t know how he killed the animal.’

(88) *Elle a peur que tu ne sois là.* (French; example taken from Zeijlstra, 2004)

‘She’s afraid you might be there.’

Furthermore, there is no evidence that *na*, *naxi* or *nantan* are clitics in the sense of French weak *ne* or affixes in the verbal complex of Bantu languages. I therefore assume

⁹⁴ Note further that the absolute negation markers in Santome are *inô*, *nô* or *não* ‘no’.

that it is a base-generated negation marker merged preceding the also base-generated TMA-markers. An immediate consequence of these findings is that the syntactic analysis given for French, where *ne* raises with the verb, fails to apply to Santome.

I will now turn to the prediction made at the start of this section, namely that Neg1 is a syntactic head. A first piece of evidence in support of the head status of *na*, *nanta* and *naxi* comes from the relation between N-words and Neg1, as in the following examples.

- (89) *Nê ũa ngê na bi fa.*
 Not one person NEG come NEG
 ‘Nobody came.’
- (90) **Personne** n’est venu. (French)
 ‘Nobody came.’

I will assume with Zanuttini (1991) and others that the NC reading reflects a specifier-head relation. This is fully in line with Zeijlstra cross-linguistic findings showing that preverbal negative particles are syntactic heads and trigger NC.

The head status is further confirmed by a number of tests presented in Merchant (2001), who discusses the head or specifier status of negation markers in a cross-linguistic perspective. In his squib, he provides an additional diagnostic for what he labels “the phrase structural status of negative markers cross-linguistically”. It is argued that only negative markers with an XP status are able to occur in the expression *why not?*, under the assumption that *why* is an XP and only maximal projections can adjoin to XP (Chomsky 1986). Thus the English negation marker *not*, claimed to be an XP, fares well in this construction. Italian, on the other hand, which exhibits an X⁰ negation marker, *non*, fails this test, requiring the use of some other negative adverb (*no*).

- (91) a. Why {**not**/***no**}?
 b. *Perchè {**no**/***non**}?

Zeijlstra (2004: 143) further claims that this test holds for all the languages on which he focused in his dissertation, except of course for those languages where XP and X⁰ negation markers cannot be set apart because of their identical status (e.g. Spanish *no*).

As Merchant points out, these facts nicely correlate with the findings of a number of other constructions, such as negative stripping and negative conditionals. The examples are drawn from Merchant (2001):

- (92) a. Anna left, but {**not/*no**} Ben.
 b. Anna é partite, ma Ben {**no/*non**}.
- (93) a. If he comes, it'll be fine; if {**not/*no**}, we have a problem.
 b. Se arriva, bene; se {**no/*non**}, avremo problemi.

Applying these syntactic tests to Santome gives the following results. In the first place, the *why not?* test itself is not available in Santome, since questioning a negative sentence requires an affirmative question:

- (94) - *Zon na ka be fa?*
 Zon NEG ASP go NEG
 'Doesn't Zon go?'
- *Punda kê kwa?*
 Why what thing
 'For what reason?' (i.e. why not?)
- **Punda na (fa)?*⁹⁵
 why no(t)

There is no restriction on the testability of the related tests proposed by Merchant, which yield the following results:

Constituent negation / contrastive negation

- (95) (**na*) *Zon ku Maya (*na)*.
 NEG Zon with Maya NEG
 ('Not Zon and Maya')

⁹⁵ Note that Neg1 cannot appear in this expression either: **punda na*.

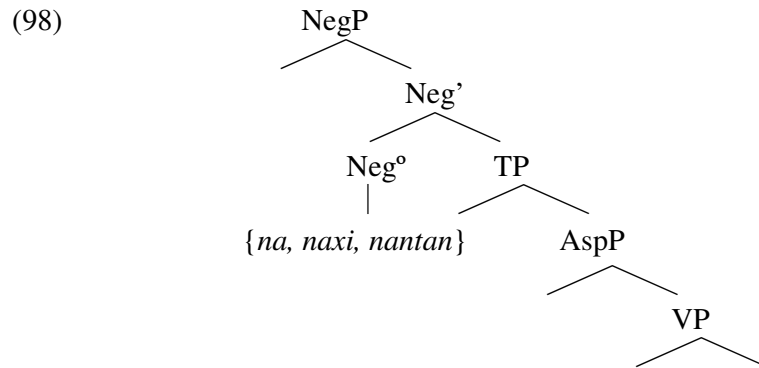
Negative stripping

- (96) *Zon be, maji (*na) Maya (*na).*
Zon go but NEG Maya NEG
(‘Zon went, but not Maya.’)

Elliptical protases of conditionals

- (97) *Da mu pa n pega jesu. *Axi na, bô ka ba kônsê ngê ku*
give 1SG for 1SG take plaster. If NEG 2SG ASP go knowperson REL
sa ami.
be 1SG
‘Give (it) to me so I can take the plaster. If not, you will get to know me.’

Therefore I conclude that Neg1 in Santome is a negative head that belongs to the class of the so-called strong preverbal negative markers (cf. Zanuttini 2001).⁹⁶ Following most literature on negation, I assume that *na*, *naxi* and *nantan* head NegP. This is illustrated in the following simplified structure.



⁹⁶ In her terminology, weak preverbal negation markers (typically clitics adjoined to V°) cannot occur by themselves, requiring the presence of an additional negative element. Although, Neg1 in Santome typically occurs together with Neg2, it was shown that this is not necessarily the case.

4.3.2. The status of Neg2

Neg1 has a clear syntactic status that complies with the general assumptions on negation markers, but what about Neg2, which is typologically rare (e.g. Kahrel 1996) and has been understudied?

In the tradition of NegP (e.g. Pollock, 1989; Ouhalla, 1990 and Haegeman, 1995), clausal negation is a Specifier-Head relation. In a language like French, this relation is assumed to be lexically visible: *ne* is the head, *pas* the specifier. In languages that lack, for instance, a lexically realized specifier, it has been argued that an empty operator occupies the specifier position. Languages whose sole negation marker is a specifier, in turn, arguably exhibit an abstract morpheme in the head position. This line of argumentation relies heavily on a universal principle that regulates negation within a clause, formalized by excellence through the Neg-Criterion, proposed by Haegeman (1995).

In more recent literature, however, a new approach to Neg2 has been sketched for a couple of languages. The basic idea of this approach is that not every language with discontinuous negation exhibits the Specifier-Head relation that has been typically assumed for French. For some languages of the Gbe cluster (Aboh 2004, *forthc.*) and in Afrikaans (Oosthuizen 1998, Molnarfi 2002, Bell 2004), it has been argued that both negative elements should be treated as heads, calling therefore into question the classic Spec-Head analysis.

The head-status of Neg2 in Santome follows from a number of facts, in particular the inability to be moved, in (99), to receive stress and to be modified by adverbs, in (100).

(99) (*Fa) Zon **na** ka kume (*fa) pixi (*fa) ku mon *(fa).
NEG Zon NEG ASP eat (NEG) fish (NEG) with hand (NEG)
'Zon doesn't eat fish with his hands.'

(100) Zon **na** ka kume pixi (*kwaji) fa.
Zon NEG ASP eat fish almost NEG

The fixed peripheral position of *fa* also follows from the fact that, unlike N-words in Santome, Neg2 can never precede Neg1 or occupy a preverbal position (101b-c), which is, for instance, possible in French infinite clauses (102).

- (101) a. *Fla ô na fla fa.*
 speak or NEG speak NEG
 ‘Speak or don’t speak.’
 b. **Fla ô na fa fla*
 c. **Fla ô fa na fla*
- (102) *Parler ou ne pas parler.*
 ‘To speak or not to speak.’

Finally, the fact that *fa* behaves like a bound morpheme provides additional support for the head analysis of this negation marker. Therefore I conclude that *fa* is a head in contexts of sentence negation.

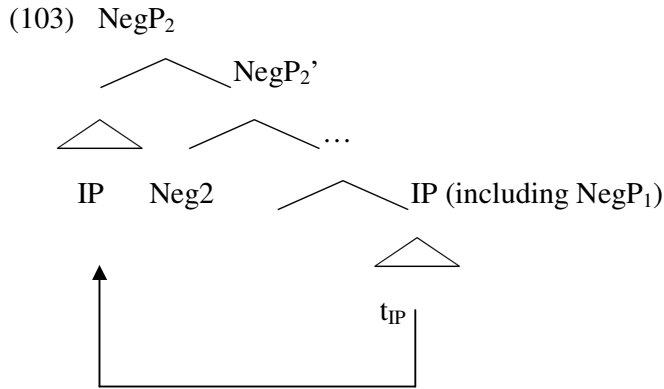
4.3.3. The relation between *na* and *fa*

In the previous sections it was shown that both negation markers in Santome are best analyzed as heads that have negative force. In standard sentence negation, these two heads typically co-occur obligatorily. The interaction between two negation markers has given rise to a number of studies that seek to disentangle in what kind of relation these markers stand to each other. The focus of these studies has been mostly syntactic. Especially since the work of Pollock (1989), the specifier-head relation between the markers occupied a central role. In a certain sense, this approach culminated with the influent Neg-Criterion, proposed by Haegeman (1995).

At that time, Kayne (1994) and Rizzi (1997) started setting the pace for a new understanding of many old questions. However controversial, antisymmetry and the fine left periphery became two important working tools in syntax that made it possible to look at negation from a different angle. Analyzing Neg2 as a head, as now proposed in some analyses for a handful of languages, opened the way for postulating a high(er), left-peripheral functional head for negation.

Albeit in different directions, Bell (2004) and Aboh (forthc.) explore these developments in syntactic theory and both propose a high basic position for Neg2, crucially higher in the structure than Neg1. Clausal movement of NegP1 and everything it contains to [Spec,NegP2] then derives the correct surface word order with Neg2 in clause-final position.

According to these analyses, Neg2 c-commands the trace of Neg1, but not the other way around.



The tree structure shows that raising IP to [Spec, NegP₂] bans c-command of NegP₂ by NegP₁, since the first node that dominates NegP₁ does not dominate NegP₂. The data of Santome, however, suggest that there is a relation of c-command between the two negation markers in which Neg₁ c-commands Neg₂. In the following example, a negative clause contains a subject relative clause:

- (104) *Maji* [_{IP} *ngê* [_{CP} *ku na sa fê fa*] [_{NegP} *na ka mendu letlatu fa*]].
 but person REL NEG be ugly NEG NEG ASP fear photo NEG
 ‘But persons that aren’t ugly do not fear photo’s.’

Neg₂ has to occur with each instance of Neg₁. Since Ross (1967), it is standard practice to analyze relative clauses as CP adjuncts. Since this adjunct is a lower embedded domain for negation, Neg₁ inside the relative clause cannot c-command the main predicate, which has therefore to be independently negated. The fact that *na* in the relative clause cannot c-command the main predicate also explains why (104) does not exhibit one single Neg₂ in sentence final position. The c-command problem in the abovementioned type of analysis can only be settled if a reconstruction analysis is adopted. However, this option is not available in Kayne’s framework and presumably not in Rizzi’s proposal as well.

It can also be shown that Neg₂ in Santome is dependent on Neg₁ and not the other way around. In the first place, it was shown that there are several situations in which Neg₂ does not occur alongside Neg₁, for instance purpose clauses or constructions embedded under a strong intensional predicate, as follows from (105) and (106) respectively.

- (105) *Lanja alkohol, bolo e, p'ê na fe flida.*
 Get alcohol rub 3SG for-3SG NEG makewound
 'Get alcohol, rub it on, so it won't cause a wound.'
- (106) a. *Bô na kunda kuma ê sa ke *(fa).*
 2SG NEG think that 3SG be house NEG
 'You don't think he is home.'
- b. *N mêsê pa bô na kunda kuma ê sa ke *(fa)*
 1SG want for 2SG NEG think that 3SG be house NEG
 'I want that you don't think he is home.'

In example (106a), the negative matrix predicate embeds a subordinate clause. Neg2 occurs obligatorily in this construction. If the matrix clause in (106a) is embedded itself under a strong intensional predicate (e.g. Quer 1998, Borgonovo 2002), as in (107b), Neg2 is precluded in the sentence final position.

Structures that exhibit exclusively Neg2, on the other hand, are typically cases of constituent negation and some sporadic instances of emphatic sentence negation, often with an emphasis particle attached to the negation marker, as illustrated in (107):

- (107) *Ni glêntu d'awa? Sabi kyê nê fô!*
 in inside of-water key fall in-3SG NEG-EMPH
 'In the water? The key didn't fall in there!'

Here, the speaker contrastively denies information in the previous discourse. An important difference between the purpose clause and the emphatic sentence above is that the grammaticality of the purpose clause relies on the absence of Neg2, whereas the insertion of Neg1 in the emphatic clause would still yield a grammatical, emphatic construction. I therefore claim for the latter case that an abstract Neg1 is present in the structure.

The presence of NegP₁ is a *sine qua non* condition for the lexical spell-out of Neg2. If NegP₁ is not projected, NegP₂ hosting *fa* is not triggered. This is shown in the following example:

- (108) *Sososo n fika sê nê ãa mina (*fa).*
 then 1SG remain without no one child NEG
 ‘Then I was left without a child.’

In (108), a negative adjunct introduced by downward entailing operator *sê* ‘without’ is able to license an N-word (*nê ãa mina*) but crucially not *fa*. Although *sê* bears a negative feature, functioning as a NC environment, *sê* does not head NegP1 and therefore Neg2 cannot be triggered.

4.3.4. Summary

The previous sections have shown that Neg1 and Neg2 are best analyzed as negative heads (NegP) in the spirit of several recent proposals. These findings alone require revisions of previous analyses with a central role for Specifier-Head relations (e.g. Pollock 1989, Haegeman & Zanuttini 1991, Haegeman 1995). I assumed that the relation between Neg1 and Neg2 relies on c-command, which is a potential problem for some analyses that will be discussed in section 4.4. Since there is nothing new or unexpected about assuming that the preverbal negation marker in Santome is a NegP that precedes the preverbal functional complex, the remainder of this chapter will primarily focus on the final element *fa*.

4.4. Previous analyses of discontinuous negation

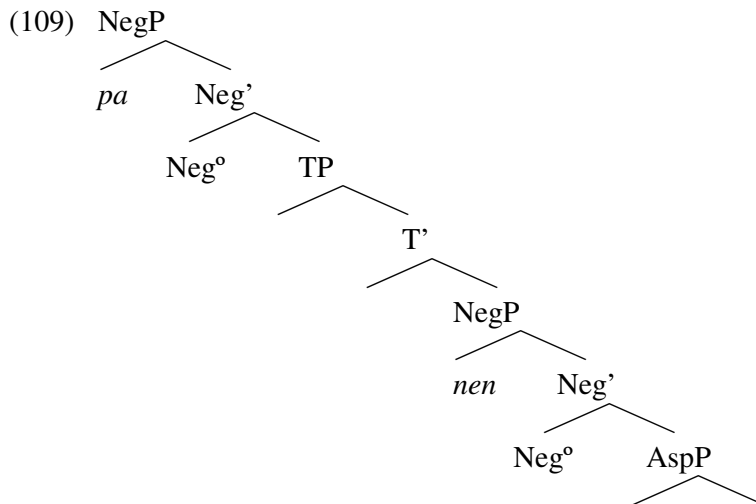
Early proposals like Pollock (1989), Ouhalla (1990), Haegeman & Zanuttini (1991) or Haegeman (1995) with a single negative projection that are able to account for a wide range of parametric variation in the expression of sentence negation are now clearly insufficient to deal with double-headed negation languages like Afrikaans or Santome. For these languages, there is an obvious quest for two different syntactic positions in order to account for the empirical data. In the following sections I will examine several analyses that account for discontinuous negation languages or which have the ability to do so. Section 4.4.1 and 4.4.2 focus on proposals that were not originally conceived for double-headed negation languages. Sections 4.4.3 and 4.4.4, on the other hand, contain proposals that account specifically for double-headed negation languages.

4.4.1. Zanuttini (1994, 1997)

The earliest claims that propose more than one projection to account for sentence negation date back to the mid-nineties. Zanuttini (1994) proposes two functional projections to encode negation: NegP and PolP. The former sits below TP and the latter dominates TP. The polarity projection PolP can be directly related to very early generativism, namely Chomsky's (1957) treatment of do-support as *Aff(irmative)*, which later became the basis of Laka's (1990) SigmaP (Σ P). In the original proposal by Zanuttini, the negation marker is generated as the head of NegP and moves to Pol^o in order to check polarity features. If [Spec,NegP] is filled, the negative specifier becomes stranded in post-verbal position, after verb movement takes place.

The PolP analysis has been adopted by several authors. Oosthuizen (1998) for instance, argues that in Afrikaans, which exhibits similarities with the negation patterns found in Santome, Neg2 heads PolP. PolP projects in the left-periphery and attracts the CP selected by PolP into [Spec,PolP]. Recently, Bell (2004) has rejected this analysis on at least two major grounds: (i) there are cases where it predicts the wrong surface structure, namely when PPs or CPs occur to the left of Neg2 and (ii) Neg2 in Afrikaans lacks an affirmative counterpart, which renders void the notion polarity. In the next section, I will discuss Bell's proposal in detail and provide data from Afrikaans.

In later work, Zanuttini (1997: 74) proposes that NegP in Piemontese can occur in several structural positions in the clause. This is illustrated in the following tree structure:



Like PolP, the highest NegP dominates TP and hosts the presuppositional negation marker (*pa*), whereas the lower NegP is dominated by TP and hosts the non-presuppositional negation marker (*nen*).⁹⁷ Moreover, the label PolP is deliberately not used in this work. Although two positions for NegP are acknowledged, they were originally conceived to account for structural variation. Yet, there are cases where *pa* and *nen* are able to co-occur, as illustrated in (110).

- (110) *Fa pa nen suli!* (Piedmontese from Lanzo; Zanuttini 1997: 75)
do NEG NEG that
‘Don’t do that!’

As Aboh (forthc.) correctly points out, the spirit of this analysis can be adjusted to double-headed negation languages by stipulating that both NegPs are obligatory in these languages and sit within the I-system (corresponding to the TMA-system of creole languages). This is the type of analysis I have adopted in Hagemeijer (forthc.) and which will be pursued in section 4.5.

4.4.2. Haegeman (2002b)

Since the work by Haegeman, especially Haegeman (1995), West Flemish has been well known for its negation patterns, illustrated in (111):

- (111) *da Valère Marie die boeken gisteren verzekest nie al gegeven (en)-eet.*
that Valère Marie those books yesterday probably not all given en-has
‘that Valère probably did not give Marie all those books yesterday.’
(West Flemish; Haegeman 1995: 157)

West Flemish exhibits optional discontinuous negation in tensed clauses, namely *nie... (en-)*, as in (111). Haegeman (1995) proposes that *nie* is a specifier of NegP that dominates TP, whereas *en* is a verbal clitic that raises with its host, V, to T. Since lexical material can intervene between *nie* and *en*, Haegeman (1995: 157) does not

⁹⁷ According to Zanuttini (1997: 67), the non-presuppositional negation marker “(...) negates a proposition without any particular discourse status”. She further claims that presuppositional negation markers sometimes behave like the default negation marker and that the non-presuppositional negation markers sometimes are presuppositional. This arguably reflects the development of the stages of Jespersen’s cycle (cf. Schwegler 1990).

pursue an analysis whereby *nie* and *en* are respectively the specifier and head of NegP. This follows from the fact that these negation markers may not be adjacent at surface structure, as can be seen in (111). In addition, Haegeman explicitly abandons the hypothesis that *ne* heads NegP and *nie* occurs in a higher functional projection, such as PolP or Σ P because this would require more high functional projections to derive the surface order.

However, Haegeman (2002b) returns to the question of higher functional projections and polarity by suggesting a PolP in the structure of West Flemish negative sentences. More specifically, Haegeman (2002b) challenges the Neg-criterion by hypothesizing that Neg^o houses an abstract operator, whereas *en-* heads PolP as a sentential negation reinforcement strategy. It follows that this gives rise to a dependency between PolP and NegP. Haegeman provides several pieces of evidence in support of the emphatic and polar function associated with *en-*. In this proposal, NegP is a lower projection that occurs in between AgroP and VP. Remnant movement is required to derive the correct surface order of tensed negatives clause with this structure. Since *ne* has lost its former role of being an independent negation marker, Haegeman suggest that this item diachronically started out as the head of this lower NegP before becoming reanalyzed as the head of PolP.

I will not adopt this type of analysis for the case of Santome, because I argued that NegP1 dominates NegP2 and not the other way around. Moreover, it was shown that Neg2 (*fa*) is not a polar element and may only bear emphasis in contrastive environments, typically in the absence of Neg1. However, emphasis is generally expressed by means of Neg2+emphatic particle, which additionally shows the neutral status of *fa*.

A virtual adaptation of PolP to the facts of Santome would be to claim that Neg1 is the head of this projection in order to account for the specific structures in which Neg1 occurs without Neg2 (cf. section 4.2.2). It was shown that many of these structures have an irrealis flavor. Yet, PolP would be expected to host other lexical material than negation, contrary to fact. The position of NegP1 in clause structure is reserved to negation. Even when Neg2 is absent, the main function of Neg1 is still to negate the clause. In light of the unmarked status of discontinuous negation, I see no advantage of adopting a label different than NegP for the projection hosting *na* and *fa*. Using the same label readily accounts for the fact that there is a relation of Agree at distance between both projections.

4.4.3. Bell (2004)

Several authors have discussed the discontinuous negation patterns in Afrikaans (Bell 2004, Donaldson 1993, Molnárfi 2002, Oosthuizen 1998; Robbers 1997), where a preverbal *nie* and a final *nie* form the default negation pattern. The distribution of final *nie* shows a number of similarities with the standard negation patterns in Santome. Example (112) shows that in matrix clause Neg2 follows adjuncts, in (113) Neg2 is blocked by the causal clause, and in (114) both the matrix and the complement clause are negated, triggering a single occurrence of Neg2 in sentence final position.

(112) *Ek het nie geslaap op die trein nie.* (Afrikaans; Robbers 1997: 35)

I have NEG slept on the train NEG
'I did not sleep on the train.'

(113) *Hy het dit nie gedoen nie omdat hy betaal is.* (Ibidem, p. 40)

he have it NEG done NEG because he paid is
'He did not do it because he was paid.'

(114) *Ek kan nie glo dat Jan nie kom nie (*nie).* (Ibidem, p. 36)

I have NEG known that Jan NEG would come NEG
'I didn't know that he wouldn't come.'

It should be noted that there are also important differences between Neg2 placement in Afrikaans and Santome. It was already mentioned that constituent negation in Afrikaans requires Neg1 and Neg2 but, more significantly, complement clauses and adjuncts PPs may occur to the right of Neg2, which is unacceptable in Santome.

(115) *Ek kan nie glo nie [dat een kind mishandel is].*

I can NEG believe NEG that one child maltreated is
'I cannot believe that one child has been maltreated.' (Ibidem, p. 35)

(116) *Sy het niks gesê nie [op die vergadering].*

she has nothing said NEG at the meeting
'She said nothing at the meeting.' (Afrikaans; Oosthuizen 1998: 88)

Oosthuizen (1998) analyzes Neg1 and Neg2 in Afrikaans as heads and proposes a high Polarity Phrase to capture the relevant data. The head of this PolP is Neg2 and the full

CP is pied-piped into [Spec,PolP]. But, as Bell (2004) points out, this analysis wrongly predicts that Neg2 is always sentence final, contrary to fact.

Bell (2004) revises and updates Oosthuizen's (1998) proposal in the spirit of Chomsky's (2001) derivation by phase. Next, I will briefly outline Bell's analysis. First, it is stipulated that there is an *Agree* relation between both negation markers whereby Neg2 dominates Neg1, i.e. Neg2 sits in a higher position than Neg1 and the correct surface order is obtained through movement operations. To make movement work, Bell (2004: 45) assumes that "*nie*₂ [=Neg2] contains uninterpretable negative features [uNeg] that are valued by the interpretable negative feature [+neg] associated with *nie*₁ [=Neg1]". The motivation for Neg2's uninterpretable features is mostly theory-internal, but he explores additional arguments within the domain of Negative Concord (NC), a feature of Afrikaans:

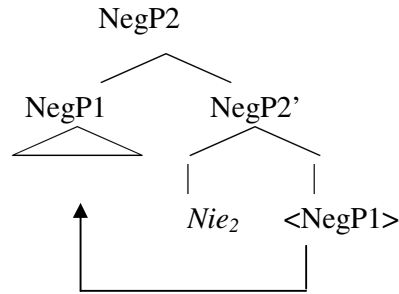
- (117) *Niemand het niks gedoen nie.* (Afrikaans, Robbers 1997:37)
nobody has nothing done NEG
'Nobody did anything.'

(Bell 2004: 46) claims that "[i]f NC is possible with multiple N-words, then we can easily justify both *nie*₁ and *nie*₂ containing interpretable features as an instance of NC." To maintain Neg2's uninterpretable features, he claims that not all dialects exhibit NC and that other languages with similar negation strategies – mostly African languages - do not have N-words at all.

Chomsky (1999) proposes two basic operations: *Agree* and *Merge*. The first operation makes use of a Probe and a Goal. Applied to the case of Afrikaans, Neg2 is the Probe, which has to eliminate its uninterpretable (OCC)⁹⁸ features against Neg1, the Goal. *Merge*, then, allows lexical movement. The following structure is a preliminary representation of Bell's proposal.

⁹⁸ OCC = a requirement that a phrase must be an occurrence of some probe and, consequently, license novel interpretations.

(118)



Since this structure is unable to account for postposed material, such as CPs and PPs, as in (115 –116) above, Bell proposes an additional XP layer between NegP2 and NegP1. In a nutshell, postposed material is scrambled to the specifier position of this XP and the material left behind is remnantly moved to [Spec,NegP2] in the fashion represented above. As a consequence, PPs and CPs are stranded to the right of Neg2, as shown by the data above. Bell argues that PP/CP-stranding is not related to information structure and often considered an afterthought, which leads to his conclusion that

“phrases and constituents stranded to the right of *nie*₂ fall outside the scope of negation. In this sense, *nie*₂ is a type of scope-delimiter, effectively marking the rightward edge of negative scope.” (Bell 2004: 141).⁹⁹

The scope-delimiting properties of Neg2 in Afrikaans are particularly clear with polarity PPs: to the left of Neg2 they enter NC and to the right of Neg2 they are ungrammatical, as illustrated in (119a) and (119b) respectively.

- (119) a. *Sy het niks gesê vir enige iemand nie.*
she has nothing said to any someone NEG
'She said nothing to anybody.'
- b. **Sy het niks gesê nie vir enige iemand.*

⁹⁹ Scope-marking, albeit in a more descriptive framework, has also been emphasized by Dieck (2000) for Palenquero. Note, however, that Dieck (2000) proposes that the discontinuous negation patterns in Palenquero are instances of right adjunction to VP and considers Neg2 a specifier. I do not adopt this type of analyses because Palenquero appears to be different from Afrikaans and Santome in the sense that Neg2 is not obligatory, although the status of Neg1...Neg2 and exclusive Neg2 structures in Palenquero is still somewhat controversial (see also Schwegler 1991).

Overall, the analysis proposed by Bell makes good predictions with respect to the placement of Neg2 in Afrikaans. Unlike previous analyses of Afrikaans, it is virtually able to capture the full range of data. Nevertheless, there are a few aspects that run into problems.

First, Bell assumes that Neg2 has uninterpretable features in the *Agree* and *Merge* framework. As he mentions himself, this is potentially problematic with respect to NC and forces him to claim that languages with bipartite negation typically do not exhibit NC, despite the fact that this property is found in several dialects of spoken Afrikaans. Now, consider the following example from Santome.

(120) *Nê ūa ngê na fe nadaxi fa.*

Not one person NEG do nothing NEG
 ‘Nobody did anything.’

This sentence exhibits an obligatory NC-reading. Hence, the argument that Neg2 has uninterpretable features because it typically occurs in languages lacking NC does not hold. This, then, undermines the assumption that Neg2 has to move and merge in order to eliminate its uninterpretable features. In other words, other arguments than the correlation between uninterpretable features and the lack of NC are required for Bell's hypothesis to work.

Second, Bell proposes an intermediate landing site (labelled XP) between NegP2 and NegP1 where PPs and CPs, or object DPs in the case of Hausa, can scramble into when they occur to the right of Neg2 at surface structure. Note also that causal clauses (*because*-clauses in Bell's terminology) have the particularity that they can occur to the right or left of Neg2 in Afrikaans. This contrast is illustrated by the following pair of sentences.

(121) a. *Hy het dit nie gedoen nie omdat hy betaal is.*

he have it NEG done NEG because he paid is

‘He did not do it, because he was paid.’ (Afrikaans, Robbers 1997: 40)

b. *Hy het dit nie gedoen omdat hy betaal is nie.*

‘He did not do it because he was paid (but for another reason).’ (Ibidem)

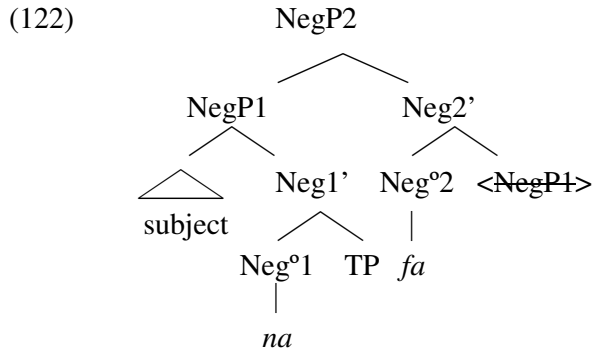
Differently from CPs and PPs, in this case a different reading obtains, which is quite obviously related to the scopal properties of Neg2. As follows from the English translation, the two interpretations are cross-linguistically common and mirror the different loci of adjunction (e.g. Lobo 2002, 2003; Haegeman 2002a) of the adjunct clause.

It is generally assumed that scrambling is either Case-driven (e.g. Russian) or information structure-related (Japanese). The first type is, of course, irrelevant for the facts observed in Afrikaans, whereas the second type is dismissed by Bell, who did not find information structure effects. However, the position of the causal clauses above with respect to Neg2 suggests otherwise. It is also tempting to treat PPs, CPs and causal clauses in a similar vein. In line with Bell's observation that PPs and CPs exhibit afterthought effects, scrambling to XP early in the derivation, before remnant movement of the remainder of vP does not seem to be a suitable solution. A different tack not explored by Bell consists of extraposing¹⁰⁰ or defocusing these phrases (cf. Zubizarreta 1998) and merging them in a high right-adjoined position, crucially outside the scope of Neg2.

Third, it is standardly assumed that Afrikaans exhibits verb-second, resembling in this respect its lexifier language Dutch. Under Bell (2004), TP immediately dominates NegP2. Thus, after NegP1 movement to [Spec,NegP2], the verb raises to T° and possibly further to C°, although this topic is not explicitly addressed. This configuration, in which TP dominates NegP, however, does not fit Santome. In the previous chapter it was shown that this language rigidly orders its preverbal TMA-material, following the order MoodP>TP>AspP. The preverbal negation marker and head of NegP, *na*, dominates these projections.¹⁰¹ Moreover, I assume with other authors working on creole languages (e.g. Veenstra 1996) that the lexicalized functional projections are base-generated. This means that Bell's analysis should be at least revised for Santome and in the following fashion. NegP2 is the topmost projection in a regular negative clause and immediately dominates NegP1, which on its turn dominates TP. Then, NegP1 is moved into [Spec,NegP2], as in the abridged representation (122).

¹⁰⁰ Oosthuizen (1998: 88) in fact suggests, for further research, that PPs to the right of Neg2 in Afrikaans are extraposed.

¹⁰¹ That NegP dominates the functional projections hosting TMA-markers is also assumed by other authors working on languages typologically similar to Santome, such as Gbe languages (e.g. Aboh 2004).



Since the verb does not move and NegP>TP, the clausal subject sits arguably in [Spec,NegP1] in Santome, whereas in Afrikaans the subject and the verb raise to [Spec,TP] and T° respectively. I consider the representation in (122) problematic because it predicts that Neg°2 scopes over the subject, meaning that *fa* is a marker that has propositional properties, and c-commands NegP1. In section 4.3.3, I argued that the presence of NegP2 is dependent on the projection of NegP1 to derive the unmarked discontinuous pattern. In the Appendix to this chapter, section 4.2, it will also be shown that *fa* is not a propositional marker like *ê*, *ô* or *fan* (cf. also data in section 4.2.1.1), which follows from its more inner syntax. Moreover, assuming that *fa* is indeed a negation marker, as I have been arguing so far, NegP2 would always exert scope over the subject. Although a subject can be focused by negation, which then becomes constituent negation (e.g. Horn 1989), it is by default outside the scope of negation, exerting scope over the predicate or any of its constituents. Therefore I will argue in section 4.5 below that *fa* is projected lower in the structure than *na*.

4.4.3.1. A short note on NEG 1 in Afrikaans

To my best knowledge, the hypothesis that there are double-headed negation languages has its origin in the research on Afrikaans by Oosthuizen (1998), whose line of research was continued by Molnarfi (2002) and Bell (2004). The projection of two NegPs in the latter author crucially relies on the X° status of Neg1 and Neg2. While I agree that the arguments concerning the X° status of Neg2 are convincing, in this section I would like to re-evaluate the arguments that support the X° status of Neg1, which sets Afrikaans apart from its closest Germanic relatives, Dutch and West-Flemish, where the negation marker is usually considered an XP (v. Haegeman 1995).

Oosthuizen (1998) shows that Neg1 in Afrikaans can be modified by certain adverbs.

- (123) *Hulle was glad nie betrokke nie.* (AFR; Oosthuizen 1998: 71)
 They were entirely NEG involved NEG
 ‘They weren’t involved at all.’

However, modification of Neg1 fails to apply in the presence of an N-Word.

- (124) a. *Dit blyk da sy absoluut niks (nie) kan onthou nie.*
 it seems that she absolutely nothing NEG can remember NEG
 ‘It seems that she can remember absolutely nothing.’ (Ibidem, p. 76)
 b. **Dit blyk da sy niks absoluut (nie) kan onthou nie.* (Ibidem)

The reasoning goes as follows. If Neg1 is the specifier (of NegP), N-word *niks* ‘nothing’ is arguably adjoined to this projection. Hence, nothing should prevent *absoluut* ‘absolutely’ in (124b) to adjoin to NegP and occur in between *niks* and *nie*. Analyzing *nie* as the head of NegP and *niks* as the specifier of this projection, the only position available for *absoluut* would be to the left of *niks*. I have no data on the position of *absoluut* in West Flemish, but it should be noted that negative constituents can co-occur with negation marker *nie*.

- (125) *Da Valère an niemand niets nie gezeid (en)-oat.*
 that Valère to nobody nothing not said *en*-had¹⁰²
 ‘That Valère had not said anything to anyone.’ (WF; Haegeman 1995: 133)

Molnarfi (2002) provides two arguments to demonstrate that Neg1 is the head of NegP. First, Neg1 cannot be topicalized, failing to trigger V2-effects. N-words are grammatical in this position.

- (126) **Nie het hy gekom nie.* (AFR; Molnarfi 2002: 231)
 NEG has he come NEG
 (127) *Nêrens voel sy veilig nie.* (AFR; Oosthuizen 1998: 62)
 nowhere feels she safe NEG
 ‘Nowhere does she feel safe.’

¹⁰² Haegeman (1995) analyzes *en* as the head of negation.

Second, in negative imperatives Neg1 (*nie*) can be moved to C, cliticized to modal *moet* ‘must’:

- (128) [CP[(*Jy*)[C *moenie*_i [NEG t_i [VP *boeke lees nie*]]]]. (AFR; Molnarfi 2002: 232)
 You must-NEG book read NEG
 ‘Don’t read books.’

The ban on topicalization is actually a contradictory argument. Since Afrikaans exhibits V2, verb movement to C° across Neg° would violate the Head Movement Constraint (Travis 1984). On the other hand, if Neg1 in Afrikaans were an XP, movement to [Spec,CP] should be allowed, contrary to fact. Note that the same restriction also applies to (Standard) Dutch *niet*, which is standardly considered an XP.

- (129) **Niet* ziet hij het. (Dutch; Zeijlstra 2004: 149)
 NEG sees he it
 ‘He does not see it.’

In the light of this restriction and a few other marginally acceptable sentences, Barbiers (2002) proposes that Dutch *niet* is lexically unspecified with respect to its categorial label. I refer the reader to Zeijlstra’s (2004: 149-151) extensive discussion of these cases, which remain inconclusive.¹⁰³ Note that in Afrikaans the co-occurrence of Neg1

¹⁰³ The only option available is to stress *niet* ‘not’ in its regular position.

- (i) *Hij ziet het NIET.*
 ‘He does NOT see it.’

However, Hans den Besten [p.c] points out that *niet* can be topicalized in certain (contrastive) environments. Note that the affirmative counterpart with *wel* can also not be fronted with the intended reading, contrary to other adverbs. With emphasis, in (iii), this is possible.

- (ii) *Hij ziet het WEL.*
 ‘He does (indeed) see it.’
 (iii) *Wel/*WEL ziet hij het.*
 ‘However, he sees it.’
 * ‘He does see it.’

Independently of their categorial label, *niet* and *wel* are more functional than other XPs. The more functional status may manifest itself in the form of syntactic restrictions such as the one in question. Haegeman (1995), for instance, considers English *not* an XP and clitic *n’t* an X°, which is motivated by their different properties. Hence, underspecification of negation markers as proposed by Barbiers (2002) should definitely stay on the research agenda.

with N-words appears to have an optional status (e.g. Molnárfi 2002, Oosthuizen 1998), which can be taken as an additional piece of evidence for Neg1's unspecified status.

As to the other case involving negative imperatives, cliticization of *nie* 'not' on modal *moet* 'must', in (128) above as it stands, is arguably be considered a case of post-syntactic phonological restructuring instead of cliticization of an X° to an X°. The following pair of declarative sentences shows that the V+NEG-movement to C° in (128) above is not the result of syntactic movement.

- (130) a. [CP *Ek* [C *moet* [XP *seker* [NegP *nie* [VP *kla* *nie*]]]]]. (Internet example)
 I must certainly NEG complainNEG
 'I certainly can't complain.'
- b. *Ek moenie toelaat dat een persoon die gesprek oorheers nie.*
 I must-NEG allow that a person the conversation dominate NEG
 'I cannot allow that someone dominates the conversation.' (Internet ex.)

The *why not?* and related tests in Merchant (2001) run into the obvious problem that Neg1 and Neg2 in Afrikaans are phonetically identical. But let us first consider the data. Note that the examples were taken from random texts on the internet. Examples (131) consists of the *why not?*-test, (132) is a case of constituent negation, (133) shows negative stripping and (134), finally tests an elliptical protasis of conditional.

- (131) *Waarom nie?*
 'Why not?'
- (132) *Nie ek nie.*
 NEG I NEG
 'Not me.'
- (133) *Ek is baie teleurgesteld vandag, maar nie verbaas nie.*
 I is very disappointed today but NEG surprised NEG
 'I'm very disappointed today, but not surprised.'
- (134) *So nie, raak 'n mens verstrik in probleme.*
 if NEG get a person tied up in problems
 'If not, a person gets tied up in problems.'

Although I do not consider these tests water-tight, these examples suggest that Neg1 may have phrasal or partly phrasal status, in the sense of Barbiers (2002). Finally, there are varieties of Afrikaans that do not exhibit NC. Since the first *nie* is postverbal, Afrikaans has thus at least two properties that *a priori* contrast with the fact that languages with preverbal negative particles are the ones that typically – but not necessarily (e.g. Bantu) - exhibit NC.

All in all, I believe that the analysis of Afrikaans as a double-headed negation language needs to be carefully assessed, since the evidence in support of the two heads is not compelling and there is evidence that suggests that one of the markers may have phrasal status.

4.4.4. Aboh (2004, *forthc.*)

Some of the languages across the Gbe cluster exhibit bipartite negation patterns similar to the ones found in Santome. Aboh (2004) and especially Aboh (*forthc.*) explore a double-headed analysis in detriment of the standard Specifier-Head approaches, couched in Rizzi's (1997) proposal for a refined functional structure in the left periphery. In a nutshell, he argues that Gbe languages lexically manifest an array of functional heads that fill the projections ForceP, TopP, FocP and FinP. Since these elements may superficially occur in the right periphery, Aboh proposes a mechanism of snowballing, i.e. successive cyclic pied-piping of chunks to the specifier of the relevant functional projections. For instance, a whole FocP dominated by TopP can be pied-piped into [Spec,TopP]. As for negation, Aboh explores two hypotheses:

- i) right-edge negation belongs to the I-system;
- ii) right-edge negation belongs to the C-system.

Hypothesis i), as Aboh points out, is a possible interpretation of Zanuttini's (1997) multiple NegPs hypothesis and was discussed in section 4.4.1. This type of analysis has been proposed in Hagemeijer (*forthc.*) and will also be pursued in more detail in section 4.5. Aboh, however, argues for an analysis that places NegP2 within the C-system, especially because of the observed interaction between Neg2 and the C-system in Fongbe. In this language, main clauses do not exhibit Neg2 but conditional clauses do, as illustrated in (135) and (136) respectively.

(135) **Kòkú má xò àsòń ń ǎ.* (Fongbe; Aboh forthc.)

Koku Neg buy crab DET NEG

(136) *Ní Kòkú má xò àsòń ń ǎ, é ná yì.* (Ibid.)

if Koku NEG buy crab DET NEG 3SG FUT go

‘If Koku did not buy the specific crab, he will leave.’

Hence, Aboh concludes that left peripheral conditions are involved in discontinuous negation in Fongbe. The technical implementation of the findings from the data follows the notion of snowballing described above. In a standard matrix clause, NegP2, the topmost projection for negation under this analysis, attracts the proposition containing NegP1 into its specifier position. This analysis therefore bears similarities with the analysis proposed by Bell, outlined in the previous section. Aboh further claims that Neg2 occupies a low position in the C-system, on the basis of example like the following:

(137) *É dù nú ǎ wè à?* (Fongbe; Lefebvre & Brousseau 2002: 135)

3SG eat thing NEG FOC INT

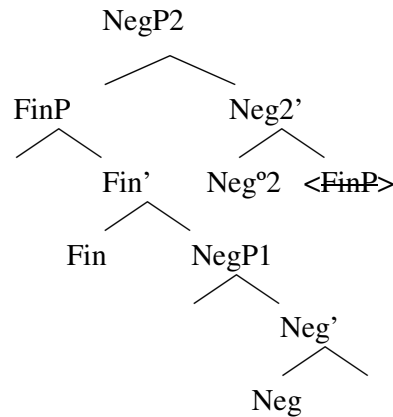
‘Is it that s/he has not eaten?’

Since the negation marker may precede other functional heads, here the Focus and the Interrogative marker, which are also claimed to be instances of the C-system, Aboh claims that NegP2 is a clause-typing functional projection sitting in the lower portions of the C-system, where it has wide scope over the proposition. More specifically, the hierarchy of functional projections proposed is the following:

(138) TopP > FocP > NegP(2)>FinP>NegP(1)>TP>...

It also important to note that in the Gbe cluster at least the following typology with respect to negation is attested, according to Aboh (forthc.):

(140)



As follows from this representation, FinP and everything that it contains is moved, or snowballed, into the specifier of the projection that immediately dominates it, namely NegP2. In this configuration, under the classical definition of c-command, NegP1 does not c-command (or scope) over NegP2.

As for the content of the final marker, Lefebvre (1998) and Lefebvre & Brousseau (2002) show that in Fongbe the negation marker *ǎ* in example (137) above is mutually exclusive with other markers that express the speakers' point of view with respect to the proposition. Aboh claims that these markers in Fongbe and Santome are evidentials. Although I have no information on the other Gbe varieties, it should be noted that Fongbe once more differs from Santome in this respect, since in the latter language negation typically precedes clause-typing markers.

I do not adhere to the view that in Santome NEg2 marks evidentiality. First, this view is not compatible with the fact that Neg2 (*fa*) is a crucial element of default negation. Second, as mentioned, it occurs in a different syntactic position than other clause-typing particles (emphasis, interrogation, etc.). Third, the definition of evidentiality doesn't apply to the contexts involving *fa*. As shown in section 4.2.2, Neg2 is only lacking in some highly specific environments, for instance certain desiderative exclamatives and purpose clauses. These domains are usually considered subtypes of epistemic modality and therefore some overlap with evidentiality, which belongs to the same domain, is not excluded *a priori*. Evidentials are concerned with the source of information and evidence type or reliability (e.g. Palmer 1986, Rooryck 2001). Typologically, notions such as direct or indirect source of information or the how the event was witnessed (visual, hearsay, etc.) are in the core of the notion evidentiality.

Crucially, none of these notions can be linked up with some of the specific negation patterns found in Santome, which are generally related to doubt and more rarely to desire or wish.

4.4.5. Summary

In the sections above I assessed the main proposals for languages that exhibit double-headed negation patterns, such as Afrikaans or certain Gbe varieties. The analyses by Bell and Aboh consider Neg2 the head of a projection, NegP2, which sits higher in the clause than NegP1. The correct surface order is derived by moving NegP1 and everything it contains into [Spec,NegP2]. In addition, some negation patterns in Afrikaans are explained by means of remnant movement.

First, the discussion of these proposals shows that languages with bipartite negation may differ to a significant extent on several aspects, such as the exact position of Neg2 with respect to clause-typing markers or the material that can occur to the right of Neg2. I argued that postulating a high NegP2 in Santome's syntax makes wrong predictions about scope and c-command relations. Moreover, I argued against the possibility that Neg2 in this creole is a C-type clause-typing morpheme, in particular an evidential marker. As a consequence, I will outline a hypothesis whereby NegP2 sits lower in the structure than NegP1, in the I-system, a possibility foreseen in the work by Zanuttini, although not meant to account for the type of double-headed negation in focus.

4.5. Towards an analysis

This section outlines an alternative analysis that accounts for the data in Santome. Differently from the analyses proposed by Bell and Aboh, I indirectly follow Zanuttini (1997) and propose a NegP2 headed by *fa* in a position in the I-system, where it is dominated by NegP1 headed by *na*.

The data discussed so far have provided a number of insights into the interaction between negation and sentence structure in Santome. The basic negation pattern was shown to consist of preverbal and clause-final negation. I have also shown evidence that Neg2 is inherently negative and behaves like a scope marker that delimits the scope of Neg1. I am further assuming that Neg1 triggers and c-commands Neg2.

Structurally, complements, both XPs and clauses, were shown to occur to the left of Neg2, even though it could be shown that Neg2 does not belong to the embedded

clause. The distribution of adjuncts with respect to Neg2 was shown to be more complex. A preliminary hypothesis consists of the intuition that adjuncts that are embraced by the discontinuous negation pattern adjoin lower in the clause than adjuncts that occur to the right of Neg2 (or in sentence-initial position). The aim of the following sections is therefore to analyze the loci of adjunction in Santome in more detail and gain deeper insights into the syntax of the projection hosting Neg2.

4.5.1. Adjunction and peripherality

As much as it is consensual in the literature that adjuncts, i.e. adverbs, adverbial expressions and adverbial clauses, affect different parts of the clause and constitute an important precision tool to determine clause structure, the theoretical approaches to adjuncts vary a great deal. In the following sections, I will focus specifically on the distribution of adjuncts with respect to negation. Within generative syntax, there are three central hypotheses with respect to the syntax of adverbs. In a nutshell, it has been claimed that adverbs are specifiers (e.g. Alexiadou 1994, Cinque 1999), that adverbs are always left-adjoined (Kayne 1994) and that adverbs adjoin to the right or the left (Ernst 2002).¹⁰⁴

I will couch the discussion of adverb placement in the Parameterized Direction Hypothesis (PDH) proposed by Ernst (2002). In addition to the arguments presented in this work, the fact that Santome is a morphologically poor language where external merge plays an important role, for instance in the case of the TMA-markers, it is conceptually counterintuitive to adopt a framework that often requires extensive phrasal movement and/or a significant number of functional projections to derive the correct linear order (e.g. Kayne 1994, Cinque 1999). Although adopting the PDH implies that adjuncts are allowed to right or left-adjoin to XP and X', I will ignore adjunction to intermediate projections for its controversial status (cf. Chomsky 1986).

In the following sections I will build on the long-standing tradition that propositional or non-propositional adjuncts exhibit different degrees of peripherality with respect to the clause. I refer the reader to Lobo (2003) for an overview. At present, I will essentially use the concept of peripherality to pinpoint the projections that host the preverbal and the final negation markers.

¹⁰⁴ For an overview and discussion of each of these proposals, I refer the reader to Costa (2004)

4.5.1.1. The distribution of adverbials

This section focuses on the distribution of certain adverbs and adverbials with respect to negation, and in particular the scope of negation. First, it should be noted that the class of monomorphemic adverbs in Santome is significantly smaller than in the lexifier language. In many cases, periphrastic constructions or items with other categorial labels, such as verbs, fill in the adverbial function. Second, the following sections will not be concerned with all the adverb types. Very high adverbials, for instance propositional modals, such as *taluvê* ‘maybe’ or *bonja* ‘fortunately, gladly’, or discourse particles like *agola* ‘well, so’ or *semple* ‘anyway’ typically have the full clause in its scope and are therefore less interesting in relation to their interaction with negation. The following sections purposely focus on three classes of adverbs whose scope often interacts with the scope of negation, namely:

- (i) focus adverbs
- (ii) time-related adjuncts
- (iii) durational adverbs

It will be shown that these commonly used types of adverbials allow us to pinpoint not only their syntactic distribution with respect to clausal functional heads, but also provide important insights into the functional projection hosting Neg2 (*fa*).

4.5.1.1.1. Focus adverbs

In Chapter 1 it was shown that Santome exhibits a restricted number of adverbs that intervene between the lexicalized functional complex and the subject and typically trigger strong subject pronouns. In addition to standard Focus marker *so* or *soku*, this group of adverbs, in the typology of König (1991), includes:

- (i) adverbs of inclusion *ten*, *tembeten*, *tudaxi*, which roughly have the meaning ‘also, too, either, as well’¹⁰⁵;

¹⁰⁵ Note that *ten* and *tembeten* may, in specific cases, function as a topic marker as well. This use is illustrated in the following examples:

- (i) *Ami tembeten ku sa ai so na sêbê nadaxi ku ska pasa fa.*
1SG also REL be here FOC NEG know nothing REL ASP happen NEG
‘As to me here, I don’t know whats going on.’

- (ii) adverbs of exclusion *tan* ‘only, just’ and *so*, which apart from being a Focus marker also has the meaning ‘only, just, very’;
- (iii) adverbs of intensification/emphasis *me* ‘-self’ and *nai* roughly corresponding ‘here’¹⁰⁶;
- (iv) scalar adverb *antê* ‘even’.

Except for *antê* ‘even’¹⁰⁷, these adverbs typically take narrow scope to their left, here the subject of the clause:

(141) *Ami me sêbê kontaji se.*
 1SG myself know tale SP
 ‘I myself know the tale in question’

(142) *Ami tan sa ai.*
 1SG only be here
 ‘It’s only me who’s here.’

When these focalizers co-occur with the standard Focus marker *so*, they obligatorily cluster to its left.

(143) *Ami tan so fika.*
 1SG only FOC remain
 ‘Only I remained.’

(ii) *Ami tembeten, xi zuxi sama mu, n ga bili funda.*
 1SG TOP if judge call me 1SG ASP open bag
 ‘As to me, if the judge calls me, I’ll open the bag.’

Not surprisingly, the left-dislocated subject sits in a projection higher than FocP, in (i), and a conditional clause, in (ii).

¹⁰⁶ In Portuguese, the best translation for *me* and *nai* would be, respectively, *mesmo* and *cá*, as in *eu cá*’ (cf. example in the previous footnote).

¹⁰⁷ Unlike the other adverbs listed, the scalar adverb *antê* has to occur to the left of the constituent it modifies.

- (i) [*Antê* [*ami*]] *ka kume pixi.*
 even 1SG ASP eat fish
 ‘Even I eat fish.’
- (ii) *N ga kume [antê [pixi]].*
 1SG ASP eat even fish
 ‘eat even fish.’

- (144) *Bô me so sa nai.*
 2SG myself FOC be here
 ‘YOU YOURSELF are here.’

Note that Focus adverbs are separated from this marker by a slight discourse break. *So* or its long form *soku* (cf. section 2.4.3) is the standard focus marker in Santome, heading a FocP that is projected on top of the NEG-TMA complex, if present in the clause, and takes scope over the material to its left.

Focus adverbs, however, are not limited to the preverbal position, since they are also able to focus constituents inside the VP, such as object DPs, adjunct DPs, or adverbs, as illustrated in (145-147) respectively.¹⁰⁸

- (145) *Ngê ka futa nganha ten.*
 people ASP steal chicken as well
 ‘People steal CHICKEN too.’
- (146) *N pali ãa mina ami tan.*
 1SG give birth a child 1SG only
 ‘I gave birth to a child by myself.’
- (147) *Sa punda ê ska tlaba bwa so.*
 be because 3SG ASP work really just
 ‘It’s because he’s working really well.’

In several respects, these examples are of course somewhat different from the cases of subject focus above. First, the expressions *ami tan* and *bwa so* in (146-147) are adjuncts and, second, the different position may change the meaning: *so*, for instance, can no longer be claimed to be the standard Focus marker. Despite these differences, so far we are exclusively dealing with cases of narrow focus, irrespective of whether the focused constituent is a DP or, for instance, an AdvP.¹⁰⁹

¹⁰⁸ Inside the VP, the nominal adverbs *nai* ‘here’ and *nala* ‘there’ have a regular locative interpretation and often behave as argumental adverbs (cf. discussion on *ba* and *be*, Chapter 2, section 2.1.4).

¹⁰⁹ There are also cases of modification of a focus adverb by another focus adverb, here with focus on the object DP:

- (i) *Ê na kume pixi tan so fa.*
 3SG NEG eat fish only just NEG
 ‘He didn’t eat exclusively fish.’

It is generally claimed that focus adverbs are maximal projections that must c-command the focused phrase (e.g. Bouma, Hendriks and Hoeksema Ms.).¹¹⁰ What seems to vary cross-linguistically are the loci of adjunction. While in English focus adverbs can attach to DPs and VPs, it has been argued that in German adjunction to arguments, either DPs or CPs, is precluded (Büring & Hartmann 2003). A first important difference between focus adverbs in, for instance, Germanic languages and Santome is the directionality of focus adverbs, which is essentially rightward in the former languages and essentially leftward in the latter. This difference also carries over to other parts of the grammar, such as Degree modification and, to a more limited extent, the DP. I assume that, like English, Santome exhibits DP-internal adjunction, which follows from the following examples.

(148) *Mosu da Zon tan pixi.*

boy give Zon only fish

‘The boy gave fish to Zon only.’

(149) *Mosu da Zon { *djandjan/*onten } pixi { djandjan/onten }.*

boy give Zon { quickly/yesterday } fish { quickly/yesterday }

‘The boy { quickly/yesterday } gave fish to Zon.’

(150) a. [*Zon tan*]_i *so mosu da* [-]_i *pixi.*

Zon only FOC boy give fish

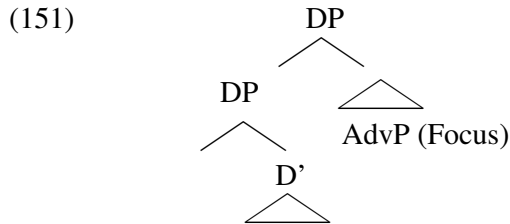
‘The boy gave fish ONLY TO ZON.’

b. *[*Zon*]_i *so mosu da* [-]_i *tan pixi.*

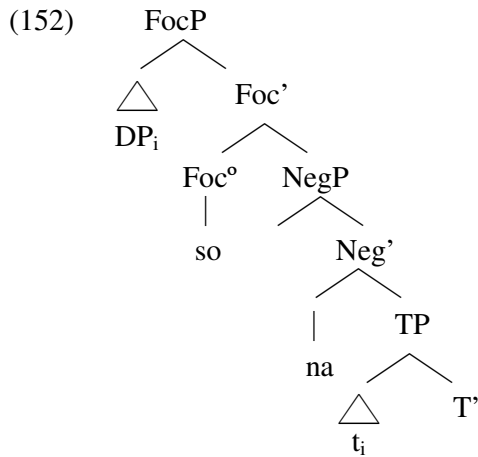
Adverb placement in double object constructions corroborates that there is a DP-internal position for Focus. In (148) it can be seen that Focus adverbs with narrow scope are able to intervene between the two objects, whereas in (149) non-DP-internal adverbs, such as manner or temporal adverbs, cannot occur in this position because they are external to the DP and therefore have to occur in VP-final position. In (150) the Goal object *Zon* has to be extracted together with the focus adverb *tan*, otherwise the sentence becomes ungrammatical, as shown by the b. example. The fact that the DP and the focus adverb have to be fronted together is another piece of evidence that this class of adverbs typically forms a constituent with the modified phrase.

¹¹⁰ However, see for instance Cinque (1999: 31), apud Bayer (1996), for the claim that these adverbs can be analyzed as heads that take their modiftees as complements.

Hence, considering the strictly local relation between most of these adverbs and the DP they modify in the cases observed above, I propose a DP-internal solution according to which these adverbs are right-adjoined to DP.



From a clause-structural point of view, this DP sits in [Spec,TP]. In the absence of higher functional projections such as NegP or FocP, no other stipulations are necessary. However, if NegP or FocP are projected, a focused DP has to raise to the specifier of the highest clausal projection. In cases like (150a), the object DP with the focus adverb would raise from its VP-internal position to [Spec,FocP].



It follows that focus adverbs used in the sense outlined above do not bear any special relation with negation apart from the fact that sentence negation only scopes over VP-internally focused material and not over focused subjects. A somewhat different picture emerges when focus adverbs take wide scope from a VP-internal position. The adverbs in question are not exclusively used for DP focus, as illustrated in (153-155).

- (153) *Mama konsê mu me.*
 mom know me really
 ‘My mom really knows me.’
- (154) *Ê tê ome ku sêbê kume tan.*
 3SG have man REL know eat only
 ‘She has a husband who is only capable of eating.’
- (155) *Ê fika ka pya mu so.*
 3SG remain ASP look me just
 ‘He just kept looking at me.’

In these examples, the final adverb scopes over the predicate. There are several ways to account for the focus adverb in this position. In theories that preclude right-adjunction as a solution for adjunct placement, it is forceful to claim that some sort of movement, e.g. VP-raising, across the adverb took place to obtain the correct surface order. Aboh (2004), for instance, argues that FocP in Gengbe is a low functional projection in the left periphery (cf. Rizzi 1997). When an item with the shape of the standard focus marker occurs in clause-final position in Gengbe, the whole sentence is moved into its specifier position. For several reasons, I believe this type of analysis is inappropriate for Santome.

First, it followed that the final position is not exclusively available for focus adverb *so*. Second, focalizers in final position can trigger different kinds of scope to their left, but *so* and *tan* can never exert scope over negation, as follows from the ungrammaticality of the c. reading in the following example.

- (156) *Zon na kume pixi tan fa.*
 Zon NEG eat fish only NEG
 a. ‘Zon doesn’t eat just fish.’ (scope over the object DP)
 b. ‘Zon doesn’t just eat fish.’ (scope over VP)
 c. ‘*Zon just doesn’t eat fish.’ (scope over negation)
- (157) *Zon na kume pixi so fa.*
 Zon NEG eat fish just NEG
 a. ‘Zon doesn’t eat just fish.’ (scope over the object DP)
 b. ‘Zon doesn’t just eat fish.’ (scope over VP)
 c. ‘*Zon just doesn’t eat fish.’ (scope over negation)

The ungrammaticality of the c. readings of these two examples shows that focus adverbs in final position cannot take scope over a full-fledged clause, which means that the whole clause cannot be moved into the specifier of a FocP in the low left periphery.

Differently from exclusive adverbs, inclusive adverbs may exert scope over material to their right, including negation:

- (158) *Zon ten na kume pixi fa.*
 Zon also NEG eat fish NEG
 ‘Zon too didn’t eat fish.’ (scope over the subject DP)
 ‘Zon didn’t eat fish as well.’ (scope over negation)

With *ten* in postverbal position, preceding Neg2, the reading with scope over negation can still be obtained, in addition to lower scope over the DP or the VP.

- (159) *Zon na kume pixi ten fa.*
 ‘Zon didn’t eat fish as well.’ (scope over the object DP or VP)
 ‘Zon also didn’t eat fish.’ (scope over negation)

In SVO languages postverbal adjuncts typically scope from the right to the left, i.e. the deeper an adjunct is embedded from a linear point of view, the higher it occurs in the structure compared to adjuncts to its left (cf. Ernst 2002: 150-2). In fact, the examples above show that the VP-final clausal adverbs under discussion can be adjoined to different syntactic positions in agreement with scope properties. In a number of cases, it might not always clear whether right-adjunction to VP or AspP and TP is targeted, but this can be tested by checking the scope of the focus adverb with respect to the only adverb that occurs within the TMA-complex, namely *kwaji* ‘almost’.

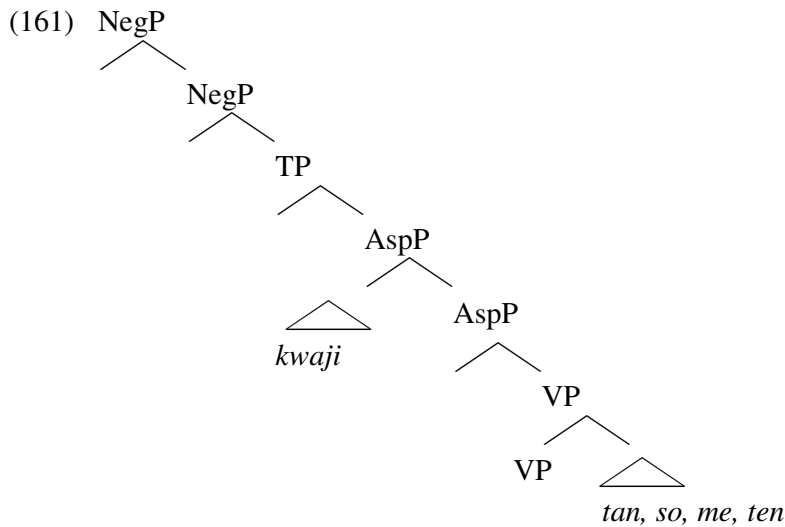
- (160) *Zon tava kwaji ka fla {so/tan}.*
 Zon TNS almost ASP speak only
 ‘Zon was almost speaking only.’
 * ‘Zon was only almost speaking.’

Hence it follows that postverbal inclusive adverbs cannot exert scope over aspectual adverb *kwaji*, which is left-adjoined to AspP. Therefore I assume that these adverbs are

VP-adjuncts. On the basis of the scope of these adverbs, it can be safely concluded that Neg2 (*fa*) has to occur higher in the structure than these exclusive focus adverbs.

Things are different in the case of the inclusive focus adverbs above. Since these adverbs are able to exert scope over negation from a VP-internal position, it is suggestive that with respect to scope they are adjoined to NegP1 (*na*) or an even higher projection but lower than Neg2 (*fa*), because otherwise these adverbs should arguably occur to the right of this negation marker, in order to preserve the correct linear order. The direct implication of such an approach would be that NegP2 sits higher in the clause than NegP1 and thus support the analyses proposed by Bell (2004) and Aboh (2004, forthc.). However, in section 4.4 I argued to the detriment of these proposals. To maintain the hypothesis that NegP2 sits in the I-system, below NegP1, I propose that inclusive adverbs are operators of predicative identity, following Matos (1992), which therefore exhibit quantifier raising at LF in order to exert scope over the extended VP.

I propose the following simplified tree structure for the position of the clausal adverbs discussed in this section:



Since adverbs do not exert scope over negation, except for the special case of *ten* (cf. also Chapter 2), it is suggestive that NegP2 is a functional projection that occurs higher in the clause than AspP. In light of the data presented in the next sections, I assume that this is indeed the case.

4.5.1.1.2. Time-related adjuncts

The title of this section is adopted from Ernst and means to include the following types of adjuncts with a temporal semantics (cf. Ernst 2002: 327). A few examples of adverbs and adverbial phrases are given for each type.

- (i) **location-time:** *za* ‘already’, *myole-wele* ‘now’, *wo* ‘agora’, *amanpasa* ‘day after tomorrow’, *sabadu* ‘on saturday’, *anu pasadu* ‘last year’, *somana ku ka bi* ‘next week’, *tê dja za* ‘days ago’, *sedu* ‘early’, etc.
- (ii) **duration:** *tudu dja* ‘always, the whole day’, *tlêxi dja* ‘for three days’, *jina nôtxi antê plaman* ‘from the night until the morning’, *jina tempu* ‘a long time ago’, etc.
- (iii) **aspectual:** *za* ‘soon’, *antawo* ‘yet, not yet’ *kwaji* ‘nearly, almost’, *lolo* ‘completely’, *kaba* ‘completely, once and for all, correctly, well, etc.’, ideophones, etc.

From a syntactic point of view, these types do not form homogeneous classes. Note also that some of these adverbs, such as *za*, may occur in more than one class in light of their meaning. Next, I will focus mostly on the scopal properties of these adjuncts with respect to negation and their possible adjunction sites.

Aspectual adverbs

Aspectuality in Santome is encoded in a number of different ways, namely by functional categories (e.g. *ka*, *ska.*), by ‘auxiliary’ verbs (e.g. *bila* ‘to turn, again’, preverbal *kaba* ‘to finish’), by morphology (e.g. *naxi* ‘not yet’, *nantan* ‘no longer, not anymore’), etc. Here I will deal with the set of items that do not fit any of these classifications and can, in that sense, be labeled adverbs. Aspectual adverbs are typically VP-final.¹¹¹

¹¹¹ Some adverbs, like *lolo* or *kaba* are strictly VP-final, whereas others, like *za* may occur in high positions, here as a topic marker.

- (i) *Oze za, inen na bi me fa.*
today already 3PL NEG come really NEG
‘As for today, they didn’t even come.’

- (162) *Sun sa ai antawo.*
 he be here still
 ‘He’s still here.’
- (163) *Ê kume non tudu kabla se lolo.*
 3SG eat 1PL every goat SP completely
 ‘He ate each and every goat of us.’
- (164) *Non ska bi b’êlê za.*
 1SG ASP come see-3SG soon
 ‘We will see him soon.’

Just like the bulk of the focusing adverbs, VP-final aspectual adverbs obligatorily precede Neg2 and negation exerts scope over the adverb.

- (165) *Wê na ska x’e antawo fa.*
 eye NEG ASP fill-3SG yet NEG
 ‘He wasn’t satisfied yet.’
- (166) *Ya ola a ka plêmê, a na ka saguji ê lolo fa.*
 PRES hour IMP ASP squeeze IMP NEG ASP shake 3SG completely NEG
 ‘When one squeezes it, one shouldn’t shake it completely.’
- (167) *Non na ka bi b’êlê za fa.*
 1PL NEG ASP come see-3SG soon NEG
 ‘We won’t see him soon.’

Note further that the complex negation markers *naxi* and *nantan* both have a counterpart with standard negation marker *na* and a VP-final adverb.

- (168) a. *Inen nantan ka pô py’e fa.*
 3PL NEG ASP canlook-3SG NEG
 b. *Inen na ka pô py’e maxi fa.*
 3PL NEG ASP can look-3SG more NEG
 ‘They can’t look at him anymore.’
- (169) a. *Inen naxi ka pô py’e fa.*
 3PL NEG ASP can look-3SG NEG

- b. *Inen na ka pô py'e antawo fa.*
 3PL NEG ASP can look-3SG yet NEG
 ‘They can’t look at him yet.’

Given the overlapping semantics of these two pairs of sentences, I assume that *maxi* and *antawo* are adjoined to a position just below negation and above AspP. Note also that, as expected, these two adverbs, as continuative and terminative VP-modifiers respectively, are mutually exclusive and exert scope over completive aspect markers such as *kaba*.¹¹²

- (170) *Inen na ka py'e { *antawo maxi/*maxi antawo } fa.*
 3PL NEG ASP look-3SG {yet/more/more yet} NEG
 (171) *Inen na ka kume kaba antawo fa.*
 3PL NEG ASP eat finish yet NEG
 ‘They haven’t yet completely finished eating.’

Prospective adverb *kwaji* (Ch. 3) does not fit the VP-final pattern observed for most of the aspectual adverbs and is exceptional in the sense that it occurs within the TMA-complex.

- (172) *Nen sa kwaji-kwaji ka xiga ke.*
 3PL be almost-almost ASP arrive house
 ‘They’re just about to arrive home.’
 (173) *Ê na sa kwaji ka nganha ke fa.*
 3SG NEG be almost ASP arrive house NEG
 ‘He’s not about to arrive home.’
 * ‘He’s about not to arrive home.’

From these examples it follows that *kwaji* precedes the aspect marker and cannot exert scope over negation. In Chapter 3 I assume this adverb to be left-adjoined to AspP. *Kwaji* does not, however, scope over VP-final *antawo*.

¹¹² Note also that completive *kaba* and *lolo* are mutually exclusive.

- (i) *Zon kume kaba lolo.
 (ii) *Zon kume lolo kaba.

(174) Zon **na** sa **kwaji** ka nganha ke **antawo fa**.

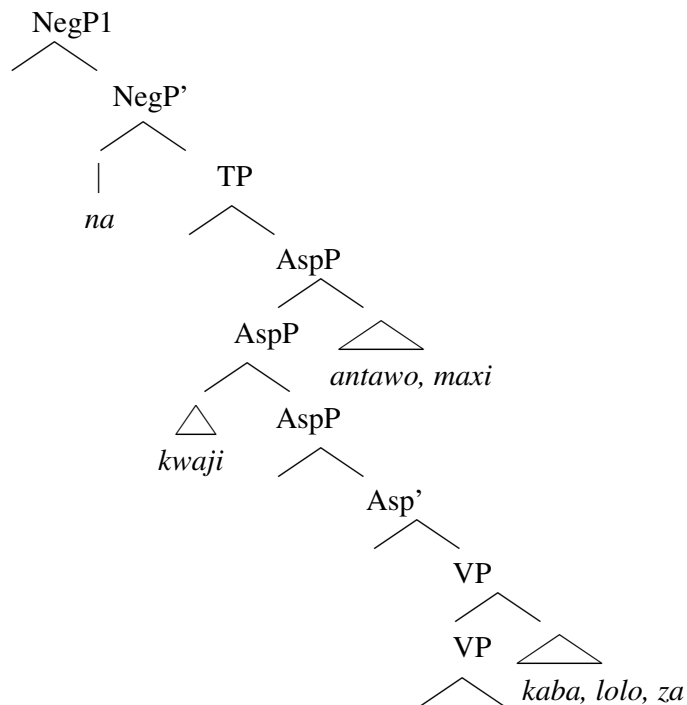
Zon isn't yet almost arriving home yet.'

From a scopal point of view, this brief incursion into aspectual adverbs with respect to negation yields the following hierarchy.

(175) *na* (negation) > *antawo* (continuative), *maxi* (terminative) > *za* (proximative) > *kwaji* (prospective) > *kaba, lolo* (completive)

It follows that negation always exerts scope over aspectual material that surfaces below negation, which is predicted from the fact that aspectual functional material occurs in the scope of negation. Even though this brief incursion shows that the order of the examined aspectual adverbs follows the order proposed in Cinque (1999) and implemented by, for instance Durrelman (2000) for Jamaican Creole, I depart away from the hypothesis that adverbs are specifiers and propose the following adjunction structure for aspectual adverbs:

(176)



4.5.1.1.3. Durational adjuncts

This type of adjuncts differs from the adverbs discussed so far in the sense that they are able to occur to the right of *fa* when they are not in the scope of negation. In the following sentences, the a. example shows the adjunct in the position where it immediately precedes Neg2 and in the b. example it follows Neg2.

- (177) a. *Zon na ka ba Lisboa [tlêxi somana] fa.*
Zon NEG ASP go Lisbon three week NEG
'Zon doesn't go to Lisbon for (a period of) three weeks.'
*'For three weeks, Zon doesn't go to Lisbon.'
- b. *Zon na ka ba Lisboa fa [tlêxi somana].*
'For three weeks, Zon doesn't go to Lisbon.'
*'Zon doesn't go to Lisbon for (a period of) three weeks.'

- (178) a. *Zon na ka tlaba [tudu plaman] fa.*
Zon NEG ASP work every morning NEG
'Zon doesn't work every morning.' (just some mornings)
* 'Every morning, Zon doesn't work.'
- b. *Zon na ka tlaba fa [tudu plaman].*
'Every morning, Zon doesn't work.'
*'Zon doesn't work every morning.' (just some mornings)

- (179) a. *Ê na ka ba kume [antê plaman] fa.*
3SG Neg1 ASP go eat until morning Neg2
'He didn't eat until the morning.'
- b. *Ê na ka ba kume fa [antê plaman].*
'He didn't eat until the morning.'

- (180) a. *A na da mu kume [jina plaman] fa.*
IMP NEG give 1SG eat since morning NEG
'They didn't provide me food since the morning.'
- b. *A na da mu kume fa [jina plaman].*
'They didn't provide me food since the morning.'

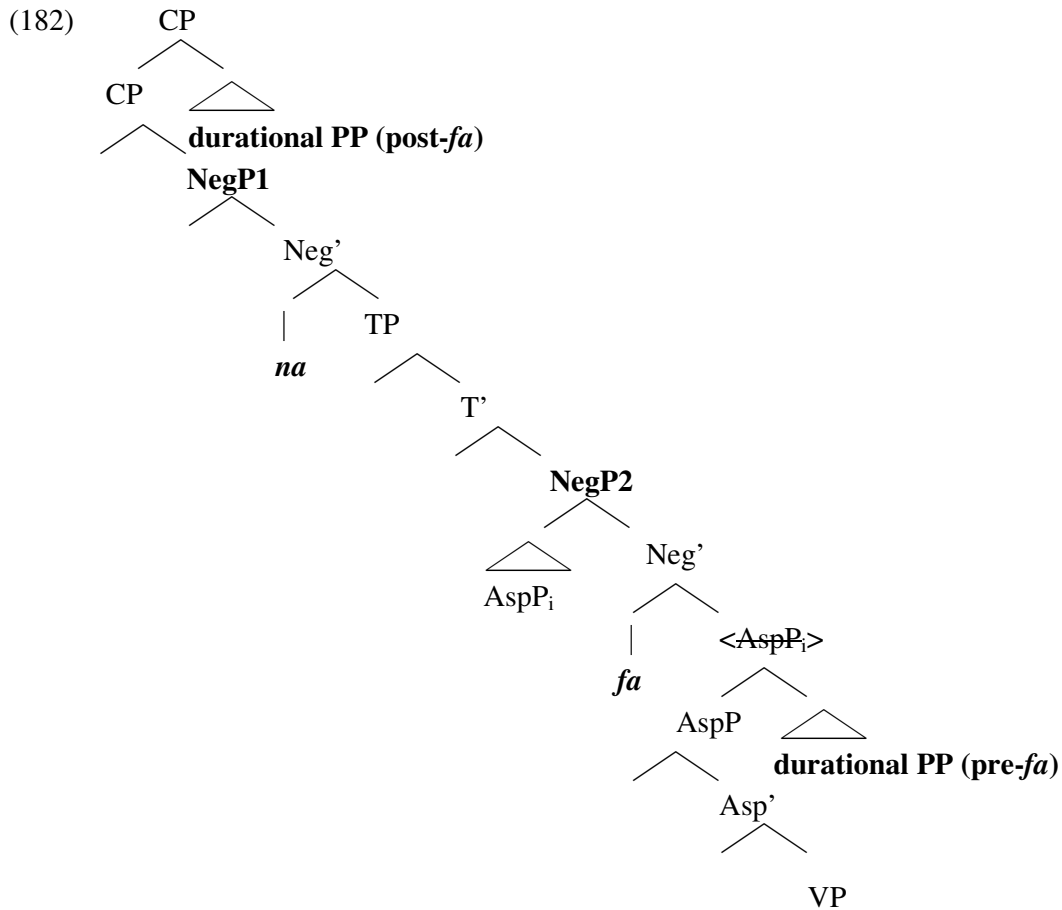
In example (177), the adjunct *tlêxi somana* ‘for three weeks’ can only be interpreted in one way according to its position to the left or the right of negation marker *fa*. The b. example would have the same reading if the adjunct occurred in pre-subject position.

Example (178) confirms these findings and shows that the quantifier *tudu* ‘every’ exerts scope over negation when the adjunct occurs to the right of *fa*. When it occurs to its left, the only available interpretation is that of negation exerting scope over the adjunct.

When a durational adjunct is headed by the prepositions *jina* and *antê*, as in (180) and (179), the difference between the pre-*fa* and post-*fa* position is not so obvious. However, it can be shown that contrastive focus can only operate on those elements that are within the scope of negation. For that reason, the adjunct *antê taji* in (181a) below can indeed be contrastively focused, whereas (181b) cannot and therefore yields an ungrammatical sentence. In (181c) the verb is in the scope of negation and can thus be under focus. This leads to the conclusion that Santome is not totally opaque to the relation between scope and focus.

- (181) a. *Ê na ka ba kume antê plaman fa, maji antê taji.*
 3SG Neg1 ASP go eat until morning Neg2 but until afternoon
 ‘He doesn’t eat until the morning but until the afternoon.’
- b. **Ê na ka ba kume fa antê plaman, maji antê taji.*
- c. *Ê na ka ba kume fa antê plaman, maji bêbê.*
 ‘He doesn’t eat until the morning but he drinks.’

In the light of the distribution of the durational phrases above, I propose the structure in (182). Note that the tree also tentatively includes the locus I propose for the functional projection hosting Neg2, *fa*, which is based on the evidence from the placement and scope of adverbs discussed in the previous sections and the findings of this section.



In this representation, durational adjuncts that pattern to the right of *fa* are merged as high right-adjuncts, arguably to CP. I do not consider NegP1 an adjunction site for these adjuncts, because in the presence of focused constituent, the adjunct cannot be stacked between FocP and NegP1, as illustrated in (183).

- (183) (*Tlêxi somana*), [*Lisboa*]_i so (**tlêxi somana*) Zon *na ka be* [-]_i *fa*.
 (three weeks) Lisbon FOC (three week) Zon NEG ASP go NEG
 ‘For three weeks, Zon doesn’t go to LISBON.’

I further assume that the same adjunct occurring to the left of *fa* is merged to AspP, which is then moved within the I-system to [Spec,NegP2], yielding the correct surface order. I will present further arguments that support AspP-raising as the basic mechanism to derive negative clauses in Santome. It will turn out that AspP-raising is a rather simple operation that is able to account for the full range of negation data, without unwarranted assumptions about c-command and scope relations.

4.5.1.2. The distribution of causal adjuncts

In the description of the negation patterns, it was shown that adjunct clauses are typically independent domains for negation, i.e. the final marker does not travel from a given clause into an adjunct clause. Causal and purpose adjuncts are an interesting test-case for scope-related properties. It is cross-linguistically common that causal and purpose adjuncts are interpreted in or outside the scope of negation in the matrix clause, as shown in the following English examples.

(184) *Marty didn't sell his bike because the gears were broken.* (Johnston 1993)

Int. 1 'Marty didn't sell his bike and the reason was that his gears were broken.'

Int. 2 'Marty did sell his bike but not because the gears were broken.'

Not only are these and some other adjunct clause types interesting with respect to the scope of negation and the syntactic implications this brings along, it is especially interesting to observe how these clauses behave with respect to Neg2 in discontinuous negation languages. The following examples are from Afrikaans:

(185) a. *Hy het dit **nie** gedoen **nie** omdat hy betaal is.*

He have it NEG done NEG because he paid is

'He did not do it, because he was paid.'

b. *Hy het dit **nie** gedoen omdat hy betaal is **nie**.* (AFR; Robbers 1997:40)

'He did not do it because he was paid (but for another reason).'

In example (185a), with Neg2 preceding the causal clause, the matrix clause is negated but the *because*-clause isn't. In, example (185b), the inverse situation can be observed. As Bell (2004) points out, the exact same reading of (185a) would be obtained by preposing the adjunct clause. Hence, in Afrikaans, the second instance of *nie* solves the ambiguity found in languages that only possess a preverbal negation marker, for instance English or European Portuguese. These facts are thus similar to the finding with respect to durational adjuncts in section 4.5.1.1.3 above. Bell analyzes *because*-clauses in Afrikaans and the other African languages surveyed in his dissertation as follows. Cases like (185b), where the adjunct clause is in the scope of negation, receive the standard analysis proposed by Bell in section 4.4.3, i.e. NegP1 and everything it contains, including the *because*-clause that is externally merged to vP, are moved to

[Spec,NegP2]. In turn, when the adjunct clause occurs to the right of Neg2, as in (185a), it is moved out of its adjunction site (vP) across lower NegP1, to a projection YP sitting between NegP2 and NegP1. Finally, to derive the correct surface order, NegP1 is remnantly moved to [Spec,NegP2]. Bell also assumes that this projection YP is distinct from the projection XP that hosts CPs and PPs that occur to the right of the final negation marker (cf. 4.4.3). This different analysis of material to the right of the final marker is motivated by the fact that the position of CPs and PPs with respect to Neg2 in Afrikaans does not trigger a change in meaning, whereas the position of *because*-clauses does have this effect.

Below I will argue against a single adjunction site for the *because*-clause, but let us first work through the facts in Santome. Causal adjuncts in this language can be introduced by *punda* ‘since, because’, *da* ‘because’ and *plôvya*, as illustrated:

- (186) *N na ka paga sapatu fa, punda n na tê djêlu fa.*
 1SG NEG ASP pay shoe NEG because 1SG NEG have money NEG
 ‘I don’t pay for the shoes, because I don’t have money.’
- (187) *Inen na fla fa da vlegonha.*
 3SG NEG speak NEG because shame
 ‘They don’t speak out of shame.’
- (188) *Zon na ka fla fa plôvya migu dê.*
 girl NEG ASP speak NEG because friend POS
 ‘The girl didn’t come because of her friend.’
- (189) *Zon na ka fla [plôvya tudu inen kwa se] fa.*
 ‘Zon doesn’t speak because of all these things.’
 Int. ‘It isn’t because of all these things that he doesn’t speak.’

Punda typically introduces clauses, whereas *da* and *plôvya* introduce DPs. In (186-187), Neg2 precedes the adjunct, which is only pattern available for *punda* and *da*, and therefore the adjunct can only be interpreted outside the scope of negation. These examples are similar to Afrikaans (185a), where the adjunct is outside the scope of negation. Example (189) is a particularly interesting case. My informants clearly prefer a reading whereby the matrix verb and the adjunct are negated and only considered other interpretations strongly marginal or ungrammatical. One of my informants even spontaneously paraphrased (189) as follows:

- (190) *Na sa plôvya tudu inen kwa se so Zon na ka fla fa.*
 NEG be because all 3PL thing SP FOC Zon NEG ASP speak NEG
 ‘It is not because of all these things that Zon doesn’t speak.’

This is somewhat surprising and differs significantly from the interpretation of the Afrikaans sentence in (185) where a wide scope reading is obtained. I therefore propose that the matrix and the adjunct can simultaneously be in the scope of negation because of the specific properties of clausal negation in Santome. Note first that this doubly-negative reading is not available in languages with two negation markers, such as French, where only the wide and narrow scope readings are available.

- (191) *Jean ne parle pas à causa de son ami*
 *Jean didn’t speak and it was not because of his friend.’

In the case of Afrikaans, where negation is apparently typologically closer to Santome, only the wide scope reading was shown to be available. I will therefore tentatively explore the hypothesis that the specific reading in Santome follows directly from the typology of negative clauses in this language and from the analysis proposed in the previous section whereby AspP is moved to [Spec,NegP].

Santome is a so-called strong or strict NC language (cf. Appendix to this chapter, section 1), where N-words can unlimitedly co-occur with both negative heads without canceling negation. I have argued that both negative heads are primarily specified for negation and that the NegPs are in a relation of Agree at distance. I propose that the fact that negation is discontinuous in syntax may also trigger discontinuous effects that follow from the syntax and semantics of each head alone. In the case of causal adjuncts that occur to the left of *fa*, this means that *na* scopes downwards over the clause in the normal fashion. However, there is another effect that derives directly from the relation between Neg2 and AspP-raising to its specifier. Hereby a Spec-Head relation is established between the material contained by AspP and negative operator *fa*. This additional property, I assume, is responsible for the double negative interpretation shown above.

This approach requires an explanation for other languages that have discontinuous negation. In the case of French, in (191), the impossibility to obtain the double negative reading may well be related with the weak negative features of *ne* and

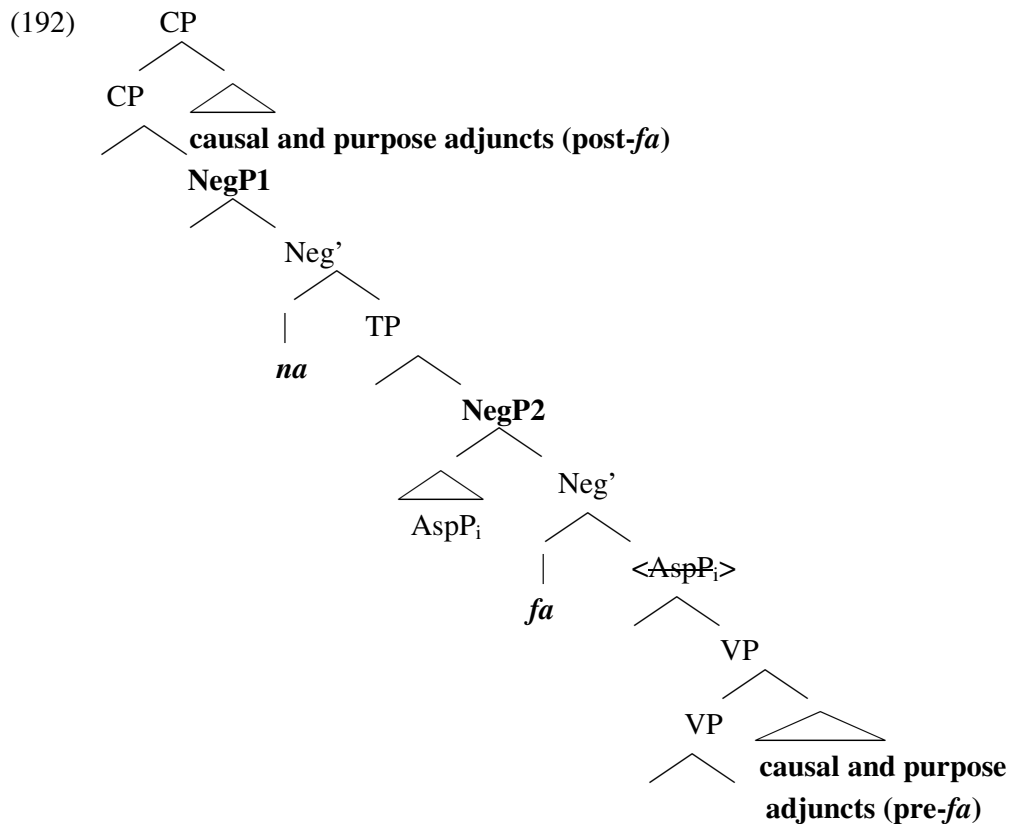
the fact that *ne...pas* is syntactically different from a language such as Santome because negation embraces the verb but not the verb and complements or adjuncts. These two reasons would lead to the correct prediction that the special double negative reading is unavailable in this language.

The difference between Santome and Afrikaans is harder to account for. I suggest that it may be related to the strength of the negative features in this language. As claimed, Santome has strong negative heads, whereas there are some reasons to believe that Afrikaans does not. Some of the data presented in Oosthuizen (1998) show that Neg2 becomes optional in the presence of polar adverbs and verbs. In fact, Van der Wouden (1994) suggests that Neg2 (*nie2*) is a NPI, which explains why it cannot occur at all in some anti-additive, monotone decreasing and antimorphic environments. According to Van der Wouden, the NPI approach may be a means to explain the somewhat idiosyncratic distribution of the negation marker.

If these assumptions are correct, the difference between Santome and Afrikaans can be seen as a difference in polarity, which may ultimately lead to a parameterized difference between double-headed negation languages: Afrikaans would project a PolP or similar (cf. Oosthuizen 1998¹¹³ and Haegeman for closely related West-Flemish) and a NegP, whereas languages such as Santome would project two NegPs. Note, for instance, that in Afrikaans constituent negation requires the presence of Neg1 and Neg2, whereas in Santome Neg2 alone negates phrases. This would be a fact compatible with Neg2's polar status in Afrikaans, i.e. its lack of an inherently strong negative feature. If Afrikaans Neg2 is indeed a special type of polarity item, the impossibility to simultaneously negate the matrix clause and the causal clause in (185) by means of a single bipartite negation structure comes for free. Even if there is a Spec-Head relation between NEg2 and pied-piped lower material, which will depend on the assumptions about clause structure in this language and on the framework adopted, Neg2 lacks independent strength to enter a strong negative checking relation. In other words, in Afrikaans Neg2 is underspecified and receives its features through its relation with NegP1. In Santome, on the other hand, Neg2 and Neg1 are fully specified and their negative features Agree at distance. In a diachronic sense, this and other facts would imply that discontinuous negation languages of the Santome and Afrikaans type may qualify distinctly with respect to the stage of Jespersen's cycle they are in.

¹¹³ Note, however, that Oosthuizen (1998) does not explore polarity in the sense of Van der Wouden (1994).

Finally, Bell claims that causal adjuncts that precede and follow Neg2 adjoin to the same projection, namely vP. The post-NegP2 position is explained by remnant movement. Instead of adopting a single locus of adjunction for causal adjuncts, I will argue that the different positioning of these adjuncts with respect to Neg2 reflects different adjunction sites, following related work by, for instance Lobo (2002, 2003) and Haegeman (2002a). Johnston (1993), in particular, argues that English *because*-clauses can be merged in two adjoined positions, VP and IP, which is similar to the analysis I will adopt. I propose that in the case of negative sentences the causal adjunct in Santome can be adjoined to VP or CP. This is reflected in the following tree structure:



The different interpretations associated to the clause when the adjunct occurs to the left of *fa* must be explained by scopal relations that involve both negative heads and the type of Agree relation that they establish.

4.5.1.3. Coordination

Another interesting and complex domain of negation in Santome are negative coordination structures, which provide further insight into the way Neg2 interacts with clause structure. The preverbal negation markers *na*, *naxi* and *nantan* (cf. section 4.3.1.1) may enter syndetic or asyndetic negative coordination. In any of these structures the subject may or not be overtly realized in the second conjunct. Note that the clauses are conjoined by *nê*, which exhibit a special polar behaviour I will discuss in more detail in section 3.1 of the Appendix. The most significant aspect of the coordination structures in (193-194) is the fact that discontinuous negation has to be repeated in each conjunct.

*Syndetic negative coordination*¹¹⁴

- (193) a. *Bô na tê mwala fa nê (bô) na tê mina fa.*
2SG NEG have woman NEGCONJ (2SG) NEG have child NEG
'You don't have a wife nor children.'
- b. *Bô naxi tê mwala fa nê (bô) naxi tê mina fa.*
'You don't have a wife nor children yet.'
- c. *Bô nantan tê mwala fa nê (bô) nantan tê mina fa.*
'You don't have a wife nor children anymore.'

Asyndetic negative coordination

- (194) a. *Bô na tê mwala fa (bô) na tê mina fa.*
'You don't have a wife nor children.'
- b. *Bô naxi tê mwala fa (bô) naxi tê mina fa.*
'You don't have a wife nor children yet.'
- c. *Bô nantan tê mwala fa (bô) nantan tê mina fa.*
'You don't have a wife nor children anymore.'

Importantly, in these structures each conjunct has to be independently negated. It is impossible to postpone *fa* until the second conjunct:

¹¹⁴ I gloss *nê* as CONJ in these examples. In section 7.1 it will be argued that *nê* is a polarity item.

- (195) a. **Bô na tê mwala (bô) na tê mina fa.*
 b. **Bô naxi tê mwala (bô) naxi tê mina fa.*
 c. **Bô nantan tê mwala (bô) nantan tê mina fa.*

Thus, negative coordination follows the restrictions observed in the data in section 4.2.1.2.2. To be sure, no arguments can be extracted out of the conjuncts, in agreement with Ross' (1967) Coordinate Structure Constraint.

- (196) a. **[Mwala]_i so bô naxi tê [-]_i fa bô naxi tê mina fa.*
 wife FOC 2SG Neg1 have Neg2 1SG Neg1 have child Neg2
 b. **[Mina]_i so bô naxi tê mwala fa bô naxi tê [-]_i fa.*

However, things are different when the coordination is established at VP-level, as illustrated in (197) and (198). Note that in these examples the conjunction *nê* is required.

- (197) *Zon na ka [da nê lêsêbê] plêsêntxi fa.*
 Zon Neg1 ASP give nor receive gift Neg2
 'Zon doesn't offer nor receive gifts.'
 (198) *Zon na tava [kloga nê kyê] fa.*
 Zon Neg1 TNS slip CONJ fall Neg2
 'Zon had not slipped nor fallen.'

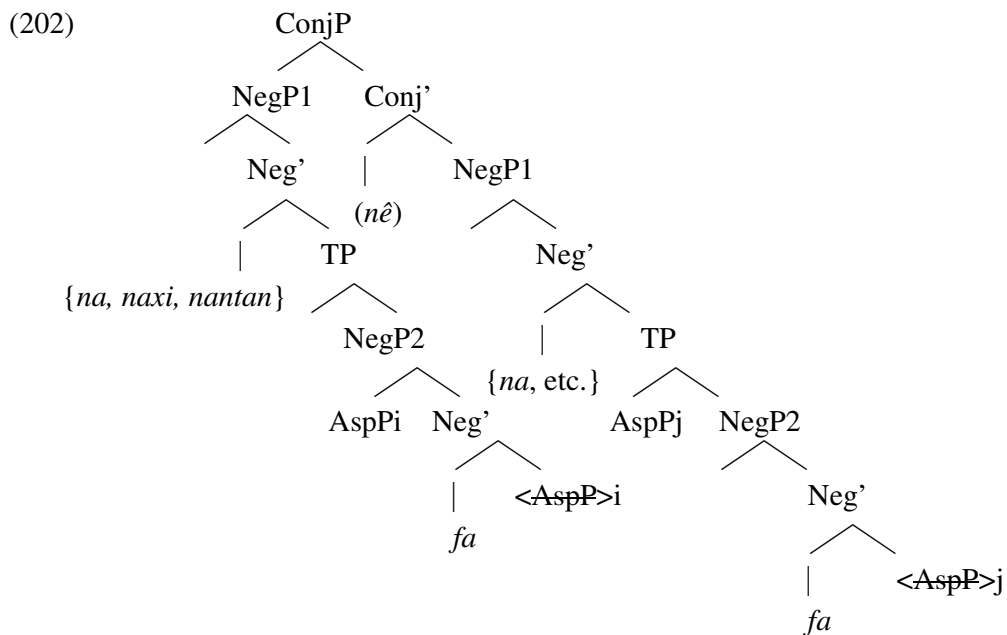
In these examples, VP-coordination requires only a single Neg2 at the end of the full sentence and not in each domain as observed for higher coordination. When the coordination is even lower, for instance at object level, an identical relation holds.

- (199) *Zon na mata [zuxi nê avogadu] fa.*
 Zon Neg1 kill judge CONJ lawyer Neg2.
 'Zon didn't kill the judge nor the lawyer.'

Furthermore, tense and aspect markers cannot establish a coordination relation.

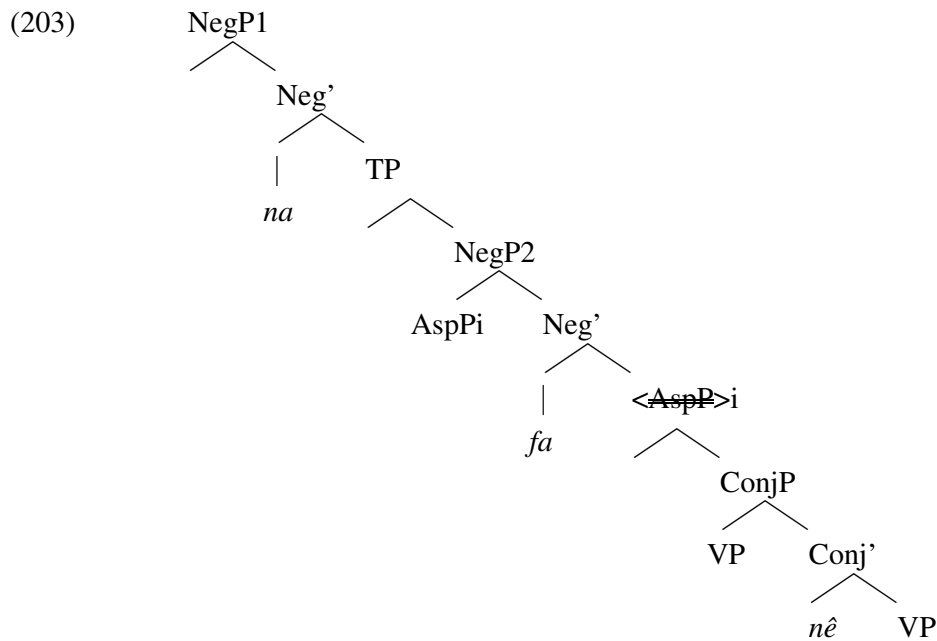
- (200) *Zon na ka da nê ka lêsêbê plêsêntxi fa.
 Zon Neg1 ASP give CONJ ASP receive gift Neg2
- (201) *Zon na tava kloga nê tava kyê fa.
 Zon Neg1 TNS slip CONJ TNS fall Neg2

As is standardly accepted, I assume that coordination can be established at different levels (IP, VP, DP, etc.) and that a ConjP projects in the syntactic structure at the relevant level (e.g. Colaço 2005). ConjP (Conj°) is headed by the conjunction or left empty in the case of asyndetic coordination. The first conjunct, the one that surfaces before the conjunction, sits in [Spec,ConjP], whereas the second conjunct is the complement of ConjP. The following reduced tree structure shows the workings of a high coordination.



This structure is coordinated by the mutually exclusive preverbal negation markers. Arguably, the subject of each conjunct is housed in [Spec,NegP]. Since each conjunct projects the I-system, including NegP2 located between TP and AspP, AspP-raising applies in the usual fashion in each conjunct. Therefore, the prediction that each conjunct has independent full-fledged negation is fulfilled. This structure also correctly predicts that extraction from each conjunct is precluded.

Coordination at a lower level in a negative sentence, for instance between VPs, is represented as follows.



Here, ConjP crucially occurs below the I-system hosting NegP2. Raising of AspP and everything it contains to [Spec,NegP2] correctly predicts that the discontinuous negation pattern projects only once, exerting scope over both VPs. This structure also correctly predicts that extraction from the VP is grammatical, as illustrated by the extraction in (204b) from the original sentence in (204a):

- (204) a. *Zon na bili poto nê fisa fa.*
 Zon NEG open door CONJ close NEG
 ‘Zon didn’t open nor close the door.’
- b. [*Poto*]_i *so Zon na bili [-]_i *nê fisa [-]_i *fa.*
 Door FOC Zon NEG open CONJ close NEG
 ‘Zon didn’t open nor close THE DOOR.’**

As follows from the indexing, extraction applies across-the-board in this example. In sum, negative coordination structures import additional evidence for the proposal involving AspP-raising.

4.5.2. Summary

The double-headed negation patterns observed in Santome and other languages show that standard analyses in which only a single head position is available (Haegeman, 1995; Laka 1994; Pollock 1989) are unable to account for the data. Under other analyses, such as PolP-NegP (Zanuttini 1994, Haegeman 2002b), negative markers or N-words are attracted out of a lower NegP to the higher functional projection PolP. However, the same problem remains, since this type of analysis departs from underlying Specifier-Head (NegP) agreement in a lower functional projection.

Recent analyses, in particular Bell (2004) and Aboh (2004, *forthc.*) adopt a double-headed analysis for bipartite negation in Afrikaans and Gbe varieties respectively. Although couched in a different framework, it was shown that in these author's analyses NegP₂ precedes NegP₁ in the syntactic structure.

Despite the mentioned advantage of double-headedness proposed in these analyses over previous analyses and the importance of this theoretical contribution to the syntax of negation, I assume the topmost NegP immediately selects TP and is co-indexed with the lower NegP, which immediately selects AspP. The correct surface order of the VP with respect to Neg₂ (fa) is then obtained by raising this AspP into the specifier of the lower NegP. Therefore, NegP₁ c-commands NegP₂ and both projections stand in an Agree relation to each other.

Following standard assumptions about clause structure, I assume that in a negative clause the subject raises out of VP and lands in the specifier of NegP (in affirmative clauses the landing site would be the specifier of TP). As argued in Chapter 3, Santome doesn't exhibit evidence for verb movement and therefore the position of both NegPs in the clause is unproblematic from this point of view.

I further assume that there is no ban on right-adjunction (cf. Ernst, 2002). Thus, modifiers like regular VP adverbs, which always follow the verb (and object), are right-adjoined items. Adopting strict linear correspondence in the sense of Kayne (1994) would unnecessarily force numerous movements that are hard to motivate given the findings, i.e. absence of verb movement and verbal morphology, externally merged lexical-functional material, etc.

The data discussed in this chapter can now be straightforwardly incorporated in the current proposal. The verb and its complements and/or complements clauses, which are selected by AspP, are raised across Neg₂ without any further stipulations and neither does the distribution of adjunct clauses with respect to Neg₂ pose any significant

problems. I argued that adjuncts can adjoin to different projections, crucially above or below the NegPs. As a consequence of AspP-raising, low adjuncts will be automatically in the scope of negation surfacing in the correct order. Adjuncts with independent tense are always high adjuncts and therefore occur outside the scope of matrix negation. In sum, it follows that discontinuous negation in Santome is another piece of evidence in support of the long-standing tradition that distinguishes peripheral from non-peripheral adjuncts.

APPENDIX TO CHAPTER 4

1. Introduction

This Appendix addresses several aspects of negation that are not nuclear for clause structure, but show that, in addition to *na...fa*, there are other promising topics related to negation. Section 2 focuses on NC and briefly discusses N-words and minimizers. In section 3, I present two cases where polarity plays a crucial role. Section 4 examines several diachronic and comparative aspects of negation in the GGC, such as the relation between the final negation marker and emphasis, Jespersen's cycle and the proto-GGC and typological features of negation in Santome in the different strata that contributed to its formation.

2. Negative Concord

Several examples in Chapter 4 indirectly showed that Santome is a Negative Concord (NC) language. In sections 1.1 and 1.2 I will provide an overview of N-words and minimizers.

2.1. N-words

Santome exhibits negative concord (NC), the presence of more than one negative element in the same clause yielding the interpretation of a single instance of negation rather than cancelling the negative interpretation. This section focuses mostly on the properties of *nadaxi* 'nothing' and *nê ãa ngê* 'nobody' (lit. not one person), but it should be noted that N-words can readily be formed by attaching *nê* 'not even' to an expression (cf. also section 2.2 on minimizers).¹¹⁵

- (1) *N naxi bêbê nê ãa tampa kaxalamba plaman se fa.*
1SG NEG drink not even one cap rum morning SP NEG
'I haven't yet drunk a single shot of rum this morning.'

¹¹⁵ Note, however, that *nê ãa* also negatively quantifies over nouns, incl pronouns, in general, as illustrated in (i) and (ii).

- (i) *Nê ãa mundu na fe mu mali fa.*
Not one world NEG do 1SG bad NEG
'Not one world has done me any harm.'
- (ii) *Punda xi bô fla nê ãa non na ka xê ai fa.*
because if 2SG speak not one 1PL NEG ASP leave here NEG
'Because if you speak, no one of us will leave this place.'

Both these items occur in subject position (2a and 3a), in object position (2b and 2b) and as objects of prepositions (2c and 3c).

- (2) a. *Nê ãa ngê na glita fô!*
 not one person NEG scream NEG-EMPH
 ‘Nobody screamed!’
- b. *Inen na tê nê ãa mina fô!*
 3PL NEG have not one child NEG-EMPH
 ‘They don’t have any children!’
- c. *Ê na ka pô kaza ku nê ãa ngê fa, sel’êlê.*
 3SG NEG ASP can marry with not one person NEG, except-3SG
 ‘She cannot marry anyone, except for him.’
- (3) a. *Sun Alê, nadaxi na pasa pa n konta sun fô.*
 Mr. King nothing NEG happenfor 1SG tell you NEG-EMPH
 ‘Sir King, nothing happened that I should tell you about.’
- b. *N na mêsê pa a f’inen nadaxi fa.*
 1SG Neg1 want for IMP do-3PL nothing Neg2
 ‘I don’t want them to do them any harm.’
- c. *N na be ku nadaxi fa.*
 1SG NEG go with nothing NEG
 ‘I went with nothing.’

In all these examples, the N-words co-occur with standard discontinuous negation, but it should be noted that certain downward entailing operators, such as *sê* ‘without’, also license N-words, as in (4). Note, however, that NC only applies within the adjunct but not between the adjunct and standard negation, as illustrated in (5). These are instances of double negation yielding a positive reading.

- (4) *Sun be dê fesa sê paga nê xintonson ê!*
 3SG go PSR party without pay not even five cents EMPH
 ‘He went to the party without paying anything!’

- (5) Ê *na be sê nê ãa kwa fa.*
 3SG NEG go without not one thing NEG
 ‘He didn’t go without anything.’ (=he went with something)

As Vallduví (1994) shows for Catalan, *without* phrases are NC environments and not cases of non-negative polarity. I will assume this is also true for Santome. It follows from the data above that Santome belongs to the class of Strict NC languages in the terminology proposed by Zeijlstra (2004). This author divides Strict NC in two types: Strict NC languages that allow for true negative imperatives (Slavic languages) and Strict NC languages that do not allow for negative imperatives (e.g. Greek and Romanian). In the relevant languages, the existence or not of negative imperatives related to inflectional morphology. Although Santome allows for negative imperatives (*Kume* ‘Eat’ / *Na kume fa* ‘Don’t eat’), the absence of inflectional morphology invalidates this criterion.

A typical feature of Strict NC languages is that N-words are frequently fronted to the initial position (as illustrated in examples (6) and (7), without affecting the single negative reading of the sentence nor absorbing standard negation (*na...fa*).

- (6) [*Nadaxi*]_i *non na tê [-]_i pa a fe kume fa.*
 nothing 1PL Neg1 have for IMP make food Neg2
 ‘We haven’t got ANYTHING to prepare food with.’
- (7) [*Nadaxi*]_i *so n na mê sê [-]_i fa.*
 nothing FOC 1SG Neg1 want Neg2
 ‘I don’t want ANYTHING.’

The obligatory co-occurrence of the standard negation marker with these N-words is the pattern found in many languages, such as Old Romance (Martins 1997, 2000), modern Rumanian (Posner 1984), Serbo-Croatian (Progovac 1994) and most creole languages, as noted by Bickerton (1981). It occurs for instance in all the Portuguese-related Atlantic creoles, including Papiamentu. NC has been particularly well studied for Haitian (e.g. DeGraff 1993, Déprez 1999).

Old Romance, and especially Old Portuguese, is an interesting starting point, because of the role it played in the formative stage of Santome. Martins (1997, 2000) shows that in Old Romance the items licensed by standard negation are weak negative

polarity items (NPI), which follows from the fact that they are often found in non-negative modal environments, a stage that is still visible in Catalan. As the result of language change in this domain, the NPI in most Old Romance varieties into strong NPI's with intrinsically negative meaning, which arguably explains why preverbal N-words suspend the occurrence of the standard negation marker in modern Portuguese and most Romance languages.

Although it seems likely that Santome inherited the general lines of this particular aspect from Old Portuguese (section 4.5.1), there are also significant differences between these two languages or between Santome and contemporary Portuguese. First, NC in Santome is virtually unbound, whereas at least in modern Portuguese it is not, as illustrated in the following examples:

(8) *N na kêlê kuma bô fla kuma Maya ba poson ku nê ãa*
 1SG NEG believe that 2SG say that Maya go town with not one
ngê fa.

person NEG

(9) *Não acreditei que tu tivesse dito que a Maria tinha ido*
 NEG believed that 2SG had-SBJV said that *the Maria had gone*
*à cidade com alguém / *ninguém.* (European Portuguese)
to-the city with somebody / nobody

'I didn't believe that you said that Mary went to town with somebody.'

This property also distinguishes French from Haitian (Déprez 1999). Second, *nadaxi* and *nê ãa ngê* do not exhibit the behavior of NPI because it can be readily shown that they bear intrinsically negative meaning in all environments. Note that this property goes back at least until the late 19th century (section 4.1) but due to the lack of early documents in Santome, it is unclear whether N-words ever went through a NPI stage.

One of the complex issues with respect to N-words is whether these items are inherently negative or not and, related hereto, whether they can be analyzed as universal quantifiers (e.g. Haegeman & Zanuttini 1991, Haegeman 1995) or indefinites (e.g. Ladusaw 1992).¹¹⁶ My sole purpose will be to show that the properties of the two N-

¹¹⁶ I refer the reader to Zeijlstra (2004) for a critical overview of these hypotheses and an alternative analysis.

words discussed in this section strongly suggest that these items are intrinsically negative.

In my corpus, the only environment where an N-word occurred without standard negation was in the presence of downward entailing operator *sê* ‘without’, as in example (4-5) above. *Sê* is arguably an operator with a strong negative feature (\neg X) and in this sense distinct from other N-word licensers of the ‘affective’ type in Strict NC languages (e.g. Bosque 1990, Giannakidou 1998). I have extensively applied typical affective environments in Portuguese and Romance to Santome, but I was not able to find a single context where N-words are licensed without clausal negation and/or have a positive interpretation.

Santome further responds positively to two classical tests proposed by Zanuttini (1991) in support of the negative quantifier status of N-words, namely modification and isolation, as illustrated in (10) and (11) respectively.

- (10) a. *Kwaxi nê ùa ngê na bi fa.*
 hardly not one person NEG come NEG
 ‘Hardly anybody came.’
- b. *Sê sêbê kwaxi nadaxi.*
 without know hardly nothing
 ‘Without hardly knowing anything.’
- (11) a. *Kêngê bi? Nê ùa ngê.*
 who come / not one person
 Who came? No one.
- b. *Kwa ku fe bô ê? Nadaxi.*
 Thing KU do 2SG EMPH / nothing
 ‘What did you do? Nothing.’

It can also be shown that isolated N-words are not necessarily discursive elliptic structures, as demonstrated in (12):

- (12) *A ka kunda bô sa ku kuzidu ni fôgô. Nadaxi.*
 IMP ASP think 2SG be with stew on fire nothing
 ‘They think you have the stew on the fire. Not at all.’

At the morphological level, it should be noticed that N-word *nadaxi* derives arguably from Portuguese *nada+assim* (lit. ‘nothing+like this’), which is exactly the same property found in universal quantifier *tudaxi* ‘everything’, from Portuguese *tudo+assim* (lit. ‘everything+like this’).

- (13) *Tudaxi n’ũa dja se me.*
 Everything on-one day SP same
 ‘Everything on the very same day.’

Furthermore, both N-words in a NC sentence can be modified by *kwaji* ‘almost’, which is hard to explain under the existential/indefinite analysis because of the single scope interpretation (cf. Zeijlstra 2004: 219).

- (14) *Kwaji nê ãa ngê na sêbê kwaji nadaxi fa.*
 almost not one person NEG know almost nothing NEG
 ‘Almost nobody doesn’t know almost anything.’

A potential argument against the hypothesis that N-words in Santome are negative quantifiers concerns universal quantifier. While the quantified DP with *yô* ‘many’ can scope over negation when it sits above negation, in (15), universal quantifier *tudu* has to be interpreted under negation, in (16)

- (15) *Yô ngê na kume pixi fa.*
 Many people NEG eat fish NEG
 ‘Many people don’t eat fish.’ (quant > neg)
- (16) *Tudu ngê na kume pixi fa.*
 all people NEG eat fish NEG
 ‘Not everybody ate fish.’ (neg > quant)

In sum, I have shown that Santome is a strong NC language, even within the domain of minimization, as will become evident in the next section. N-words in this language were shown to be inherently negative and exhibit a number of properties that favour the negative quantifier analysis.

2.2. Minimizers

In addition to the N-words above, Santome also displays a number of items in negative clauses that reinforce negation. These items have been labeled minimizers in the literature (e.g. Horn 1989, Schwegler 1990). In many languages with or without NC, minimizers are typically weak NPIs because they have no intrinsic negative meaning (e.g. ‘a red cent’ in ‘I didn’t have a red cent’).

In Santome, minimizers are always complex items with the shape *nê* ‘not even’ + Noun. I am aware of the following items: *nê sombla* (lit. not-even shadow), *nê pwêla* (lit. not-even dust), *nê pikina* (lit. not-even little), *nê uku~niuku~niku* (lit. not-even dirt), *nê minge* (lit. not-even crumb), but it is likely that more of these minimizers exist. These forms all share the meaning ‘(nothing) at all’ and their etymology can in most cases be traced back to Portuguese. As in other languages, minimizers are more restricted in their use than N-words in the sense that they are dependent on the semantic properties of the predicate they occur with (e.g. I don’t have a red cent vs. *I didn’t eat a red cent) and, probably due to this marked nature, there is some variability in the grammatical judgments of native speakers. As example (17) suggests, I found minimizers to be particularly common in the domain of eating and drinking.

- (17) *Ê na kume {nê pikina/nê pwêla/nê minge} fa.*
 3SG Neg1 eat MIN Neg2
 ‘He didn’t eat anything at all.’

In the first place, it can be observed that the internal structure of these items is identical to *nê ûa ngê* ‘nobody’. I am not aware of any minimizer that lacks *nê*, which constitutes a difference with Portuguese, where some of these items can be optionally preceded by *nem* ‘not even’, as in (18-19). Only in a few cases is *nem* obligatory, as shown in (20)

- (18) *Não gastei (nem) um tusto.* (European Portuguese)
 not spent not even one dime
 ‘I didn’t spend a dime.’
- (19) *Não levantei (nem) um dedo.* (Ibidem)
 not lifted not even one finger
 ‘I didn’t lift a finger.’

- (20) *Não vou lá *(nem) morto.* (Ibidem)
 not go there not even dead
 ‘No way will I go there.’

Like the N-words in the previous section, these items can for instance be fronted and occur in isolation, which is related to the fact that the structure of these items is identical to *nê ãa ngê* ‘nobody’. The case of minimizers in Portuguese is identical to that of Catalan (Vallduví 1994), who argues that in this language minimizers preceded by *ni* ‘not even’ are N-words, whereas the same items without *ni* exhibit the properties of polarity items. Vallduví proposes four diagnostic tests to underscore this difference:

- (i) ability to occur in isolation;
- (ii) ability to be modified by *almost* or *absolutely*;
- (iii) grammaticality in preverbal position;
- (iv) ability to appear in *yes/no* and *if* contexts with a nonnegative value;

In Catalan, minimizers and N-words respond positively to diagnostic (i-iii). As for diagnostic (iv), N-words and NPIs pattern alike, whereas minimizers are not allowed in the contexts mentioned. It was already shown in the previous section that diagnostic (i-iii) apply to N-words in Santome. Diagnostic (iv) has to be abandoned because in Santome, as well as in Portuguese, these contexts do not exhibit the polarity effect found in Catalan. The following examples correspond respectively to diagnostic (i-iii).

- (21) *Kê kwa ku ê kume? Niuku!*
 Whatthing KU 3SG eat / MIN
 ‘What did he eat? Not the slightest bit.’
- (22) *Kwaji niuku.*
 ‘Almost nothing.’
- (23) *Niuku ê na kume fa.*
 MIN 3SG NEG eat NEG
 ‘He didn’t eat the slightest bit.’

Therefore it follows that Santome behaves like Catalan. Since minimizers occur invariably with *nê* in this language, they bear inherent negative meaning and must be classified as N-words.

3. Negative polarity and underspecification

Despite the fact that Santome is a Strict NC language in the sense of Zeijlstra (2004), as shown in the previous section, there are a number of items that have exhibit a ‘special’ behavior with respect to clausal negation. In section 3.1 I will discuss coordinative conjunction *nê* ‘nor, and, (not) even’ and in section 3.2 adverb *antawo* ‘(not) yet, still’.

3.1. *Nê* ‘nor, and, (not) even’

It was already shown that *nê* can be used as conjunction in coordinated sentences (Ch. 4, section 4.5.1.3).¹¹⁷ First, *nê* may coordinate clauses or DPs, as follows from (24) and (25) respectively:

- (24) *Ome na jera fa nê mwala na pali fa.*
 Man NEG reproduce NEG CONJ woman NEG give birth NEG
 ‘The man didn’t reproduce and the woman didn’t give birth.’
- (25) *Zon na mata nê zuxi nê avogadu fa.*
 Zon NEG kill nor judge nor lawyer NEG.
 ‘Zon didn’t kill the judge nor the lawyer.’

Although at first sight the use of *nê* resembles the use of Portuguese *nem* in these constructions, it will be shown that things are quite different. I will show that simple *nê* or correlative *nê...nê* are polarity items. The first indication that *nê* might not be simply a negative conjunction like English ‘nor’ comes from the fact that Santome lacks a clause-level coordinative conjunction corresponding exclusively to ‘and’, although sometimes the borrowed form *y* ‘and’, from Portuguese *e* ‘and’, is used.

Note first that correlative negation of the type *nê...nê* can occur to the left of clausal negation, for instances as subjects, in (26), or extracted arguments, in (27).

¹¹⁷ This morpheme can be traced back to Portuguese *nem* ‘not, not even, nor’ (cf. Matos 2003). This negation marker has several functions and as clausal negation marker occurs typically - but not exclusively - in coordination (*não...nem* ‘not...nor’; *nem...nem* ‘nor...nor’). In modern Portuguese, this item does not exhibit polarity. Like Portuguese *sem* ‘without’, *nem* is negatively specified. Santome *sê* is derived from *sem*. Unlike *nê*, Santome *sê* bears is fully specified for the negative function.

- (26) *Nê zuxi nê avogadu na pô blaga demanda se fa.*
 CONJ judge CONJ lawyer NEG can solve quest SP NEG
 ‘Neither the judge nor the lawyer are able to solve this quest.
 (Quintas da Graça, 1989: 19)
- (27) [*Nê matabisu nê lomosu*]_i *manu na bi da mu [-]_i fa.*
 CONJ breakfast CONJ lunch brother NEG come give 1SG NEG
 ‘Not breakfast nor lunch did he bring me.’

In these cases, the clauses have a standard NC reading, similar to what was shown for N-words in section 2 above.

However, there is an important contrast with sentences that lack standard clausal negation. Crucially, in these cases *nê* does not have a negative meaning, as follows from instances of correlative coordination (28-29) and simple coordination (30-31).

- (28) *Nê kompa nê sun se tê sotxi.*
 CONJ godfather CONJ 3SG SP have luck
 ‘The godfather and the man are lucky.’
- (29) *Y mina se ku ka nansê ku nê êlê nê men dê*
 and child SP REL ASP be born REL CONJ 3SG CONJ mother POS
ka môlê?
 ASP die
 ‘And what to say about a child and its mother that die at birth?’
- (30) *Yô-yô ngê, nê ministru, ku a mata.*
 many-many people even minister that IMP kill
 ‘Many, many people, even ministers, were killed.’
- (31) *Êlê so sa kapataji di tudu ngê ku sa vivu nê ku sa motxi.*
 3SG FOC be foreman of all people REL be alive and REL be dead.
 ‘He’s the foreman of all people that are alive and dead.’

Example (28-29) are cases of coordinated subjects, example (30) shows the inclusive function of *nê*, and in (31) *nê* coordinates two relative clauses.

The following pair shows that *ne...nê* is in fact a polar item. After an affirmative question, in (32), correlative coordination receives an affirmative reading, whereas a negative question, in example (33), triggers a negative reading.

- (32) *Kêngê ka ba ple? Nê êlê nê mwala dê.*
 who ASP go beach /CONJ 3SG CONJ wife POS
 ‘Who went to the beach? He and his wife.’
- (33) *Kêngê na ka ba ple fa? Nê êlê nê mwala dê.*
 ‘Who didn’t go to the beach? Not him nor his wife.’

These constructions presumably involve ellipsis. After a negative question, NegP is recovered in the elided structure, yielding a negative correlative coordination. In the other case, NegP does not project in the answer because there is no evidence for this projection in the question and therefore the negative reading is not available.

Hence, I conclude that *nê* exhibits polarity sensitivity, acquiring its negative meaning under the scope of negation. It follows that *nê* is in fact a polarity-sensitive item, although it does not comply to the standard definition of NPI in the sense that these items are only licensed and grammatical in negative contexts, but crucially not in affirmative environments (e.g. Hoeksema 2000).

3.2. Antawo ‘(not) yet’

Another clear-cut case of polarity concerns the adverb *antawo* ‘still, yet, not yet’. Examples (34-36) show the affirmative meaning of this adverb.

- (34) *Sela bô mwe pikina antawo fa.*
 must 2SG learn little still EMPH
 ‘You still have to learn a little more about housekeeping.’
- (35) *Ê na ka môlê fa. Ê ka bila skapa antawo.*
 3SG NEG ASP die NEG / 3SG ASP turn escape again
 ‘He doesn’t die. He will escape again.’
- (36) *Ê sa mina pikina? Efan, antawo.*
 3SG be girl small / yes still
 ‘Is she still a virgin/young girl? Yes, she still is.’

In or around negative environments, however, *antawo* acquires a negative meaning, as illustrated in (37) and (38).

- (37) *So nen na bi antawo fa.*
 then 3PL NEG come yet NEG
 ‘Then they haven’t come yet.’
- (38) *Sa bô naxi tê daji di netu san antawo fa.*
 well 2SG NEG have age of grandson POS yet NEG
 ‘Well, you don’t have her grandson’s age yet.’

Antawo cannot be modified, but it follows respectively from examples (39) and (40) that *antawo* occurs in isolation with a negative meaning and that it may sit higher than negation.

- (39) *Oze Sabino na tê dwentxi fa? Dwentxi mu? Di flega? Antawo.*
 Today Sabino NEG have patient NEG patient POS to rub not yet
 ‘Today you don’t have any patients? Patients of mine? To rub? Not yet.’
- (40) *Antawo sun na mêsê fa?*
 still you NEG want NEG
 ‘You still don’t want it?’

Antawo is a polarity item whose properties are identical to the correlative coordination *nê...nê* in the previous section.

4. Diachronic and comparative perspectives

This section focuses on the origins and diachrony of the negation patterns in the GGC. I will first present older data of negation in Santome and discuss the relation between negation and emphasis. Second, I will provide a brief overview of negation patterns in the other GGC and propose a historical reconstruction of standard negation in these languages. Finally, I will place the negation patterns in the GGC in a typological perspective, comparing several dominant features with the different strata that contributed to the formation of the GGC.

4.1. 19th century sources

As mentioned in the previous chapters, the earliest attested examples of sentence negation in Santome date back to the late 19th century and already show the discontinuous patterns we find in the contemporary language. The source that contains

most examples of negative sentences is Negreiros (1895), but a few examples can also be found in earlier work by Schuchardt (1882) and Coelho (1880). In fact, in light of the data presented, it is possible that these authors had either partially common sources or knew each other's work. Most of the proverbs in Coelho (1880) are found in Negreiros (1895) as well. Coelho provides the following negative sentences:¹¹⁸

(41) *Fingui mole, fingui nan mole fan; auá sugá, nomi na cabá fan.*

mouse die mouse NEG die NEG water drie name NEG finish NEG

'The mouse died or didn't die; the water dried, the name didn't end.'

(42) *Glavana pô sa longo, fia guinhon ni bódó d'auá na cá sugá fan.*

dry season may be long leaf watercress on bottom of-water NEG ASP

sugá fan.

dry NEG

'Even though the dry season may last, the leave of the watercress at the bottom of the river doesn't dry.'

In another section of his work, Coelho (1880) provides the following examples with the emphatic negation marker *fô*.

(43) *Nó bendê claçon fô.* (=N na bëndê klason fô)

'I didn't sell trousers.'

(44) *No sabê fô.* (=N na sêbê fô)

'I don't know.'

Schuchardt (1882) discusses negation only very briefly and mentions, *en passant*, that similar to French *pas* or *point* Santome exhibits *fá* or *fó*, suggesting that *fá* may be related to Portuguese *fava* 'fava bean', a hypothesis I have discussed in Hagemeyer (2003). The only example provided in his work is the following, with the mention that the final marker can be *fá* or *fó*:

¹¹⁸ Note that throughout this section I have preserved the original orthography and the glossing and translations are of my own responsibility.

- (45) *Un ne sabê fá / fô.* (Schuchardt 1882: 914)
 1SG NEG know NEG
 ‘I don’t know.’

Negreiros (1895) presents a significant number of negative sentences, most of them proverbs:

- (46) *Ninguê ná mole ni mundu si dêfêtu fan.*
 nobody NEG die in world without bad habits NEG
 ‘Nobody dies in this world without bad habits?’ (Negreiros 1895: 342)

This example also shows that negative quantifiers in subject position co-occur with the standard negation pattern. Moreover, examples (47-49) respectively show the use of emphatic *fô* and preverbal negation markers *naxi* and *nantan*. In (49) it can be seen that the subject relative clause is independently negated, a property that is also found in the modern language (cf. Ch. 4, section 4.3.3)

- (47) *Nã tê qué fô.* (Negreiros 1895: 353)
 1SG-NEG have house NEG-EMPH
 ‘I have no house.’

- (48) *Cume qu’n cumê zá ná tam buá dá mun di cume fã.*
 Food REL-1SG eat alreadyNEG good for 1SG to eat NEG
 ‘The food that I’ve already eaten is not good to me; it has already become sour.’
 (Negreiros 1895: 162)

- (49) *Mina muála cu náchi bôá fa, ná cá tandji homê ni liba*
 young woman REL NEG good NEG NEG ASP entertain man on top
pêma fã. (Negreiros 1895: 340)
 palm tree NEG
 ‘The girl who hasn’t become a woman yet doesn’t entertain a man in a palm tree.’

The following table sums up the graphic representation of sentence negation markers used by the authors above.

Table 1. Graphic representation of negation markers in Santome in the 19th century.

Negation	Coelho (1880)	Schuchardt (1882)	Negreiros (1895)
Preverbal marker	<i>na, nan, no, nó</i>	<i>ne</i>	<i>ná, nan, na, nã, náchi, ná tam</i>
Final marker	<i>fan, fô</i>	<i>fá, fô</i>	<i>fâ, fã, fan, fãn, fô</i>

Despite the lack of homogeneity, the table and the data show that discontinuous negation was certainly a feature of the late 19th century's language. It also follows that the preverbal tripartite system (*na, naxi, nantan*) and the emphatic and neutral final marker (*fô* and *fa*) already co-existed.

The graphic representation of the negation markers is also a point of interest. The forms *no, nó* and *ne* in Coelho and Schuchardt are presumably misrepresentations of sentence negation, but it is possible that the information at their disposal concerned the absolute negation marker *nô* or *inô* 'no'. In this sense, Negreiros (1895) should be considered the most reliable source.¹¹⁹ This being settled, it follows that the locus of variation particularly concerns the nasality of Neg1 and Neg2. I will return to this issue in section 2.2.

The only relevant 19th century source that provides empirical information about the syntactic distribution of Neg2 is Negreiros. In fact, the data in his work show that there no significant differences with respect to Neg2's placement in the 19th century and in contemporary Santome. *Fa* cannot reach into causal subordinate clauses, conditionals or polarity adversative clauses, which all form independent domains for negation, in (50), (51) and (52) respectively.

(50) *Plôcu ná piá ó sé fan, [pundá ê çá zudê].*

pig NEG look at heaven NEG because 3SG be jew

'The pig doesn't look at the sky because he's a Jew.' (Negreiros, 1895: 336)

¹¹⁹ Note also that the pamphlets written by Francisco Bonfim, in the early 20ties and in the middle of the 20th century systematically use *na/naxi/nantan...fa* for sentence negation.

- (51) [*Chi bô cá nã çá butchizádu fa*], ‘ngá butchizá bô cu
 if 2SG MOOD NEG be baptized NEG 1SG-ASP baptize 2SG with
*nômi de Padê, de Filhu, de Splitu Santu.*¹²⁰
 name of Father, of Sun of Holy Spirit
 ‘If you weren’t baptized, I would baptize you in the name of the Father, the Sun
 and the Holy Spirit.’ (Ibidem, p. 352)
- (52) *Béga chiá ná cá blócê ninguê fã, cé lá fómi*
 stomach full NEG ASP upset person NEG only hunger
 ‘A full stomach doesn’t bother anybody, hunger does.’ (Ibidem, p. 349)

Following the analysis in Chapter 4, I assume that the structures in (50-52) are cases of peripheral adjunction. Non-peripheral adjunct clauses, on the other hand, are within the scope of negation as in the contemporary language, as in (53).

- (53) *Á ná cá vôlô cu mina [cé lumiá men dê] fã.*
 IMP NEG ASP insult with daughter without name mother POS NEG
 ‘One doesn’t insult the daughter without offending the mother.’ (Ibidem, p. 349)

Finally, there are several instances of contexts without Neg2 corresponding to the irrealis type, namely negation in a relative clause, in a hypothetical environment and in a *pa*-construction (cf. Ch. 4, section 4.2.2).

- (54) *Padê cu ná tê môçu cá clágá missáli bá glêsa.* (Ibidem, p. 338)
 priest REL NEG have boy ASP carry missal go church
 ‘A priest without an altar boy carries the missal to the church.’
- (55) *Iá pá sun mun ná scá gáná mun* (Ibidem, p. 354)
 PRES for 2SG NEG ASP cheat 1SG
 ‘It seems you are cheating on me.’

Proverbs are an interesting source of information, because they often go back to other stages of the languages and, due to their particular status, are less likely to be affected by grammatical change. Daio (2002) provides an extensive list of 790 proverbs, a few

¹²⁰ Note also that mood marker *ká* (*cá*) precedes the negation marker in this example (cf. section 3.3.8.1). Given the lack of other examples, it is not clear whether this reflects a mistake or an older pattern.

of which show that negated relative clauses do not always exhibit Neg2, as in (54-55) above.

- (56) *Ngê ku na tê sotxi na ka toma kodo fe lasu fa.*
people that NEG have luck NEG ASP take rope make knot NEG
'People that are unlucky don't take a rope and make a knot.' (Daio 2002: 74)

The same goes for hypothetical environments.

- (57) *Xi moska na kuji da flida (fa), kê kwa ku ê ka kume?*
if fly NEG answer to wound (NEG) what thing KU 3SG ASP eat
'If the fly doesn't respond to a wound, what will he eat?' (Daio 2002: 22)

Since both relative clauses and conditionals can be subsumed under the irrealis type, the possibility that negation in these constructions shifted historically from an exclusive preverbal pattern (Neg1 only) to the default pattern that requires both Neg1 and Neg2 is suggestive.

Finally, I was unable to find any instances of Neg2 as the sole negation marker for contrastive purposes (Ch. 4, section 4.2.3). In conclusion, with respect to sentence negation, no significant changes have occurred in little more than a century, even though some data suggest that negation in relative clauses and conditionals may have shifted from an exclusive preverbal pattern towards a default discontinuous pattern.

4.2. Negation and emphasis

In the previous section it was shown that Neg2 in Negreiros (1895), and also in Coelho (1880), is frequently represented as nasalized *fã* or *fan*. This form also exists in contemporary Santome but is used in emphatic affirmative environments. In example (58), *fan* emphasizes an imperative clause, in (59) it contradicts negative prior discourse, and in (60) it emphasizes a constituent.

- (58) *Fô wê mu fan!*
get out eye POS EMPH
'Get out of my sight!'

(59) *Bô na ka fe ôtlo kwa fô?*
 2SG NEG ASP do other thing NEG-EMPH

‘Don’t you do anything else?’

Ami?N ga fe fan.

1SG 1SG ASP do EMPH

‘Me? I sure do.’

(60) *Ami me fan, mina mu!*

1SG myself EMPH child POS

‘I myself, my child!’

However, this marker also takes on the form *fa*, which was first observed by Ferraz, and thus becomes homophonous with Neg2.

(61) *Ba non fa.* (Ferraz 1979:118)

go we **EMPH**

‘Please let us go.’ / ‘Shall we go?’

(62) *Aglasa mu sa Ernestino fa.*

Name POS be Ernestino **EMPH**

‘My name is Ernestino.’

Ferraz (1979) refers to this *fa* as a respectful morpheme, but this is not confirmed by the data in my corpus, which show that *fa~fan* is actually used as a generalized emphasis marker. Some speakers consider that the emphasis marker has to be always *fan*, but that this form may have very weak nasalization. Note further that the emphatic negation marker *fô* cannot be used for affirmative emphasis. Like other emphatic particles, such as *ô* in (63), *fan* is also able to follow the negation marker (64) and may occur in interrogative environments (65).

(63) *Sun na-a tôlô fa ô!*

He NEG-be silly NEG EMPH

‘He (formal) is not silly!’

- (64) *Na kêsê non fa ê.*
 NEG forget 1PL NEG EMPH
 ‘Don’t forget us!’
- (65) *Kyê avo, punda Dêsu avo, na da mu fa fan!*
 EXCL grandma because God grandma NEG give me NEG EMPH
 ‘Oh, please grandma, don’t beat me!’
- (66) *Sun dêsê ku pêtú fan?*
 you go down with chest EMPH
 ‘Do you climb down with your chest [against the tree].’

However, I found sporadic cases where *fan* occurs in the position reserved to the negation marker:

- (67) *Sangê na-a plopi men bô fan.*
 lady NEG-be proper mother POS NEG-EMPH
 ‘She’s not your real mother.’
- (68) *Sun Alê, kidalê, na mata padjin mu fan ê.*
 Mr.King please NEG kill godfather POS NEG EMPH
 ‘King, please, don’t kill my godfather.’

Example (68) is particularly interesting for the present purpose. Discontinuous negation is obligatory in imperative clauses, which means that *fan* functions as a negation marker. This is confirmed by the fact that *fan* is followed by emphasis markers *ê* and *ô*. Thus, Santome is not the type of language where emphasis or insistence markers associate with a standard negation marker but do not have negative content. Languages with this typology are, for instance, Krongo (Kahrel 1996) or Fongbe (Lefebvre & Brousseau 2002). That *fan* in (68) cannot be claimed to be an insistence marker also follows from the fact that two of these markers are mutually exclusive:

- (69) *Bô sa tólô {ê / ô /*ê ê /*ê ô/... }!*
 2SG are silly EMPH
 ‘You are silly!’

The ungrammaticality in (69) can be explained by the fact that it would be intuitively awkward to express different attitudes towards a single proposition.¹²¹ Hence, the findings show that Neg2 in Santome is inherently negative and occurs in a specific structural position. Except for the examples above, where *fan* is part of the discontinuous negation pattern, affirmative *fa(n)* cannot be preceded or followed by other discourse particles, suggesting it occupies the same position and is subject to the restriction that one cannot express different attitudes at the same time.

I have not found emphasis marker *fan* in the other GGC. This form is arguably the result of the contraction of *fa* and discourse particle *an*, which has essentially interrogative and emphatic functions. As an interrogative particle, it normally occurs in clause-final position (70), although it may also occur in clause-initial position (71). As an emphatic particle, it is particularly productive with speech verbs such as *fla* ‘to speak’, as in (72), where it has the function of drawing attention to what the speaker is going to say next.

(70) *Kê mina di Adon ku fe mu kwa se an?*

What child of Adam REL do 1SG thing SP INT

‘What child [of Adam] did that to me?’

(71) *An, bô bi tlabà?*

INT 2SG come work

‘Did you come to work?’

(72) *Mosu ê, santome fla an: pobli na ka lomosa fa,*
 boy VOC Santome speak EMPH poor NEG ASP lunch NEG
pobli ka kume.

poor ASP eat

‘Hey boy, in Santome we say: the poor don’t lunch, they eat.’

Not surprisingly, there are also instances of *an* co-occurring with Neg2.

(73) *Ê xê ni ke p'ê nantan bi fa an.*

3SG leave from house for-3SG NEG come NEG EMPH

‘He left home to never come back.’

¹²¹ For discussion of these markers (in Haitian Creole), see for instance Lefebvre (1998: 213-217).

Note further that the prediction is borne out that *an* cannot co-occur with other emphatic particles, such as *ê*, *ô* or *fan* for the reason pointed out above.¹²²

In sum, the examples in this section show that, although negative *fa* and affirmative *fa(n)* usually behave distinctively from a syntactic and semantic point of view, the historical data and a few pieces of data from contemporary Santome suggest that emphatic *fa~fan* and negative *fa* functionalized from a common source. Assuming that this is indeed the case, there are basically two hypotheses that can account for the data:

Hypothesis A: the final marker started out as an intensifier of negation (e.g. *fava*, Schuchardt 1882: 914) and specialized for affirmative environments as well.

Hypothesis B: the final marker started out as an affirmative discourse particle and specialized for negation as well.

There are reasons to believe that Hypothesis A better accounts for some of the facts. In the first place, all the GGC exhibit a similar final negation marker, as will be shown in section 4.4, Table 2, but only Santome and Fa d'Ambô exhibit the intensifying function. Second, in section 4.5.3 it will be argued that the Neg2's syntactic patterns are related to the distribution of Neg2 in Kongo languages. In these languages, Neg2 arguably had its origin in emphatic or intensifying particles. Thus, a possible scenario is one in which the semantics and syntax of a single item with both negative and emphatic functions was borrowed into the proto-creole. This would be the corollary of many other specific overlapping facts. That a form corresponding to emphatic *fan* is not found in Lung'ie and Ngola may perhaps be explained by the development of an exclusive final negation marker, a path that was completed in Lung'ie but not in Ngola. This, then, may have prevented a single clause-final item with two interpretations, negation and emphasis, from co-existing in these languages.

¹²² There is an exception to this claim, namely the chants that are embedded in folk stories, for instance:

(i) *Abi ô, abi ô, Toni ê, Toni ô an. Awa ê, Toni ê ô ô an.*

In these cases, the expressiveness of particles is fully exploited. These strategies never occur in normal speech.

4.3. Negation in the Gulf of Guinea creoles

One of the many grammatical domains where the genetic relation between the GGC is fully transparent is negation. In this section I will briefly discuss negation patterns in Lung'ie, Ngola and Fa d'Ambô and propose a historical reconstitution of negation in the proto-GGC.

4.3.1. Lung'ie

This language differs from the other three GGC in that it exhibits almost exclusive final negation. This feature dates back at least as far as Schuchardt (1889), who provides the following sentence.

- (74) *Un mé sê fá.* (Schuchardt 1889: 469)
1SG want NEG
'I don't want (it).'

Schuchardt's findings are confirmed by Günther's (1973) grammar of Lung'ie. The following example shows that *fa* also occurs in final position, but to the left of emphatic particles, and that Lung'ie exhibits NC.¹²³

- (75) *na vída mé n sa pódí fá da nígé-ᵿ nómí ifí péné sé, ki*
in life POS 1SG ASP can NEG to nobody name hair SP REL
n té na upétu fá ó!! (Lung'ie; Günther 1973: 122)
1SG have on chest NEG EMPH
'I swear by my life that I can't tell the names of the three hairs that I have on my chest!'

Crucially, Günther claims that the negation was originally discontinuous, with *na...fa*, and that this pattern can still very sporadically be observed. Moreover, there are still cases where *na* occurs without *fa*, namely final clauses:

¹²³ The glossings are of my own responsibility, since Günther did not gloss the stories he collected

- (76) *pa či na kudá ...* (Lung'ie; Günther 1973: 78)
 for 2SG Neg1 think
 'so that you don't think ...'

Although *fa* in Lung'ie is subject to similar constraints as Neg2 in Santome with respect to the scope of negation, its syntax differs from Santome in some respects. As Maurer (forthc.) points out, *fa* occurs to the left of complement clauses when the matrix verb is negated.

- (77) *n sebé fa ó!, Ji a pagá, sũ mÉ!* (LU; Günther 1973: 130)
 1SG know NEG EMPH if IMP pay Sir
 'I don't know if they paid, Sir.'

Günther's stories also exhibit sentences that are semantically negated only once but where negation marker *fa* occurs twice. The examples involve a relative clause and a serial verb construction, in (78) and (79) respectively.

- (78) *a sebé máli fa ki e fezé fa.* (LU; Günther 1973: 114)
 IMP know evil NEG REL 3SG do NEG
 'One doesn't know the evil he did.'

- (79) *a sa pódì mará 'li fa pwé idètu isólu sé sódudí fa*
 IMP ASP can tie up 3SG NEG put in sun SP without blaim NEG
 'One can't simply tie him up and put him in the sun without blaim.'

Since none of the other three GGC exhibit Neg2 in these positions, this suggests that the absence of a preverbal negation marker required *fa* to shift towards the negated predicate.

4.3.2. Fa d'Ambô

The earliest sources on Fa d'Ambô, Vila (1891)¹²⁴ and Barrena (1957)¹²⁵, show that this language exhibits the standard discontinuous negation pattern *na...f*. However, the

¹²⁴ Previously to Vila's publication, Schuchardt published Vila's notes in *Kreolische Studien VII* (1888).

¹²⁵ Barrena's (1957) grammar was published posthumously. He died in 1925 (Post 1997).

paper on negation in Fa d'Ambô by Post (1997) provides many details of which I will only highlight a few for the present purpose.

First, the negation patterns in Fa d'Ambô and in Santome are very similar. Neg2 takes syntactic scope over relatives, complement clauses and *pa*-clauses but not over causal clauses, for instance. Unlike Santome, however, Fa d'Ambô exhibits constructions, namely relative and purpose clauses, where final negation marker *f* occurs twice while semantically there is only a single instance of negation.

- (80) *Amu na sa sini-f pa amu kubili xadyi-f.* (FA; Post 1997: 297)
 1SG NEG be zinc-NEG for 1SG cover house-NEG
 'I do not have zinc to cover my roof.'

Post assumes that *f* is a clause-final clitic, which is arguably the case in all the GGC. Differently from the other three GGC, however, the descriptions suggest that Fa d'Ambô lacks NC. Where the other GGC use the NC strategy, Fa d'Ambô exhibits negative polarity, as illustrated in the following examples:

- (81) *Bo na da zuguán ngue-f.* (FA; Barrena 1957: 44)
 2SG NEG give some person-NEG
 'Don't give it to anybody.'
- (82) *Zuguá ngue na cunji-f?* (Ibidem)
 some person NEG answer-NEG
 'Nobody answers?'

According to Barrena, the indefinite quantifier *zuguá* 'some' is used for persons and *zuguán* 'some' for things. It was shown that N-words in Santome occur without *fa* (*nê ãã ngê* (**fa*) 'nobody'), because they are inherently negative. As expected, in Fa d'Ambô the negation marker is required to provide a negative reading.

- (83) a. *Zuguá ngue-f.* (FA; Barrena 1957: 34)
 some person-NEG
 'Nobody.'

- b. *Zuguán ja-f.* (Ibidem)
 some thing-NEG
 ‘Nothing.’

Similar to Santome, Neg2 negates constituents and *f* or *fa* are also found in emphatic affirmatives.¹²⁶

- (84) *Ku bo-f.* (FA; Post 1997: 305)
 with 2SG-NEG
 ‘Not with you.’
- (85) *Se sa xosy-fá!* (Ibidem, p. 307)
 CONJ be thing-EMPH

Fa also occurs in interrogative sentences, although Post isn’t conclusive about whether it has a negative value.

- (86) *Se ma-se fa: o-Tusantu Magavu-fa?* (FA; Post 1997: 308)
 CONJ sir-DEM say EMPH-Tusantu Magavu-NEG
 ‘And the man said: isn’t that Tusantu Magavu?’

Like Lung’ie and Santome, negative purpose clauses typically occur without Neg2:

- (87) *Osyi ku eli pe olemu sa pa batelu na fo buka.* (FA; Post 1997: 308)
 when that3SG put paddle be for canoe NEG go turn
 ‘When he puts the paddle, it is to avoid that the canoe turns.’

Post concludes that Neg2 is a scope-related clause final clitic.

4.3.3. Ngola

Maurer (1995) shows that Ngola also exhibits discontinuous negation, but that often Neg1, *na*, is reduced to *a* or simply lacking. The final marker *wa* is clearly distinct from

¹²⁶ In her conclusions, Post (1997: 314) questions whether the clause final negation marker is just homophonous with emphatic *fa* or whether we are dealing with a generalized emphasize that occurs with negation as well.

the other three GGC and appears to be in free variation with *va* and *fô*. Maurer considers the latter form a borrowing from Santome, albeit without emphasis associated. Despite the different phonetic shape of the final markers, Neg2 typically occurs in a similar final position.

- (88) *Nê ãa no na thêka bê ôtô wa.* (NG; Maurer 1995: 131)
 Not one 1PL NEG ASP see other NEG
 ‘No one of us sees the other.’

This example also shows that Ngola exhibits NC, but it should be noted that instances of polarity can also be found.

- (89) Q: *Kwai bô mêthê?* (Ibidem, p. 133)
 thing 2SG want
 ‘What do you want?’
 A: *Kwa wa.*
 Thing NEG
 ‘Nothing.’

Note that this answer could also be the periphrastic N-word *nê ãa kwa* ‘nothing’ (lit: not one thing), without negation marker *wa*. Furthermore, like in the other GGC, purpose clauses lack Neg2.

- (90) *Ê lôkê kwa bisi rê pa ê na nana.*
 3SG clean.up thing dress POS for 3SG NEG spoil
 ‘He cleaned up his clothes so they wouldn’t spoil.’ (Ibidem, p. 132)

4.4. Reconstructing negation in the GGC

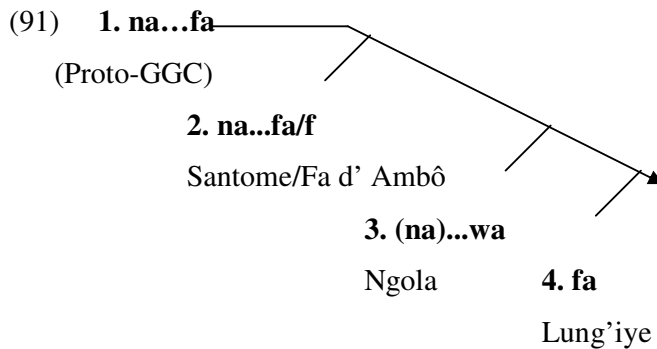
The following table illustrates the basic negation patterns in the four GGC and whether these languages exhibit an affirmative emphatic particle that overlaps with the final negation marker.

Table 2. Negation patterns in the GGC.

	Negation				Affirmation
	default negation	purpose clauses	constituent negation	emphatic negation	emphasis
Santome	<i>na...fa</i>	<i>na...</i>	<i>...fa</i>	<i>na...fô</i>	<i>fa~fan</i>
Fa d'ambô	<i>na...f</i>	<i>na...</i>	<i>...f</i>	<i>?na ... fa</i>	<i>f~fa</i>
Lung'ie	<i>...fa</i>	<i>na...</i>	<i>...f</i>	<i>...fa ô ~ f'ô</i>	-
Ngola	<i>(a~na)...</i> <i>wa ~ va ~ fô</i>	<i>na...</i>	<i>...wa</i>	<i>?wa ê</i>	-

In addition to the identical phonetic form of the negation markers in the table, there is a significant amount of overlap between the syntactic position of Neg1 and Neg2 in all the GGC. Neg1 typically occurs between the subject position and the TMA-material, whereas Neg2 typically marks the end of the VP and reaches into certain subordinate domains, such as complement and relative clauses, but not into causal or conditional domains, for instance. Purpose clauses are another outstanding shared property, since the four GGC all exhibit a pattern with exclusive preverbal negation, even Lung'ie, whose standard negation pattern consists of the final marker alone.

According to Table 2, Santome and Fa d'Ambô are arguably the most conservative GGC with respect to negation, whereas Lung'ie is the most innovative creole. The similarities between the negation patterns in the GGC uncontroversially show that they developed from a common early source, a proto-GGC (e.g. Ferraz 1979, Hagemeyer 1999, Schang 2000). Hence, considering the contemporary patterns, I claim that a discontinuous negation pattern (*na...fa*) can be reconstructed for the four GGC and that this pattern existed prior to diffusion in time and space. The following diagram illustrates the proposed scenario.



4.5. Typology and transfer

In the following sections it will be shown that Santome, and arguably the other GGC as well, preserved different features of negation found in the main strata that contributed to the formation of these creoles.

4.5.1. Portuguese

Etymologically, preverbal *na*, attested in all the GGC, can be traced back to Portuguese *não*, which follows, for instance from the nasalized complex negation marker *nantan* ‘no longer, not anymore’ and from the nasalized form *nã* used by Negreiros (1895). The syntax of the preverbal negation marker, however, is not necessarily exclusively related to Portuguese. In Edoid, the Nigerian language cluster that presumably imposed a founder effect on the GGC, negation also occurs between the subject and the TMA-markers (e.g. Agheyisi 1990, Omoruyi 1989). In the western Bantu cluster, negation can typically either precede or follow non-topicalized subjects (Güldemann 1996) and can therefore also not be discarded as possible input. Thus, rather than assigning the preverbal negation marker’s syntax to Portuguese, it is preferable to assign a multi-source or conflated origin.

A feature that can arguably assigned to Old Portuguese, however, is NC. In Martins (1997, 2000), it is shown that in Portuguese N-words went through a stage where they behaved as weak polarity items, which is, for instance, visible from the fact that negative subjects were able to co-occur with the standard negation marker, in (92), which is obligatory in Santome (93).

- (92) *Nenhuu nom mostrava que era faminto.* (Old PTG; Martins 2000: 194)
No-one NEG showed that was starving
 ‘Nobody showed that he/she was starving.’

- (93) *Nê ãa ngê na têndê glita fa.*
 not one person NEG hear scream NEG
 ‘Nobody heard the screaming.’

However, I argued in section 2 that at least in contemporary Santome N-words have intrinsic negative features and therefore differ from weak polarity items. In Old Portuguese, the standard negation marker eventually started to drop, a tendency correlated to the shift from weak to strong polarity items.

- (94) *Nenhum (*nãõ) mostrava que era faminto.* (Contemporary PTG)
 ‘Nobody showed that he/she was starving.’

Another feature that is arguably related to Portuguese is NC with minimizers or intensifiers.

- (95) *Ê na kume nê minge fa.*
 3SG NEGeat not-even crumb NEG
 ‘He didn’t eat anything at all.’

- (96) *Nãõ vale isso nem migalha.*
 NEG is-worth this not-even crumb
 ‘This ain’t worth a crumb.’ (PTG; Gil Vicente, *Farsa dos Almocreves*)

In fact, in the work of Gil Vicente, minimizers appear to occur typically with *nem*, for instance *nem figo* ‘nothing’ (lit. ‘not even a fig’) or *nem palha* ‘nothing’ (lit. ‘not even straw’). In Vicente’s play *Romagem dos agravados*, there is even an instance of a contraction of the negative marker and the minimizer, namely *nemigalha* ‘nothing’ (lit. not even a crumb’), instead of *nem migalha* in (96). In contemporary Portuguese, like in Catalan (Vallduví 1994), minimizers can generally occur with *nem*, as shown in section 2.2, whereas *nê* in Santome was shown to be compulsory. I have no information on the status of *nem* in older stages of Portuguese.

The link with Portuguese in this submodule of grammar follows also from the fact that languages of the Edoid cluster and the relevant Bantu languages do not exhibit NC nor the type of intensifying strategy sketched above.

4.5.2. Edoid

Under the hypothesis that Edo or languages from the Edoid cluster were the dominant founder languages, which follows from historical and linguistic evidence, what imprint, if any, did these languages leave on the negation patterns? The available descriptions on these languages show that they do not exhibit a final negation marker (e.g. Omoruyi 1989, Agheyisi 1990, 1991, Melzian 1942), which is confirmed by Ben Elugbe [p.c.] and Ronald Schaefer [p.c.].¹²⁷ It follows that negation in the Edoid cluster is preverbal and differs in several respects from preverbal negation in Santome. First, Edo itself exhibits two distinct negation markers, *ma* and *i* (Agheyisi 1991, Omoruyi 1989).¹²⁸ Second, these markers carry tense-aspect information, *ma* provides past tense and *i* non-past.

- (97) *Òzọ má rì èvàré.* (Edo, Agheyisi 1991: 14)
Ozo **NEG-PAST** eat food
'Ozo did not eat food.'
- (98) *Òzọ í rì èvàré.* (Ibidem)
Ozo **NEG-NON-PAST** eat food
'Ozo does not eat food.' / 'Ozo will not eat food.'

Nevertheless, Edo and Santome arguably share the same tripartite preverbal morphosyntax of negation, as illustrated in Table 3. I have also included the strategies used in the other GGC, which show that Ngola and presumably also Fa d'Ambô use similar strategies.

¹²⁷ It should be noted, however, that according to Elugbe [p.c.] Ghotuo, a north-central Edoid language, exhibits a discontinuous negation pattern that is realized by means of tone. The historical implications of this pattern have not been studied, as far as I am aware.

¹²⁸ Omoruyi (1989) also mentions the negation marker *ghé* for imperatives but, as he points out, this is arguably the contraction of an auxiliary and negation marker *i*.

Table 3. Preverbal negation markers in Edo and Santome.

	<i>not</i>	<i>not anymore/no longer</i>	<i>not yet</i>
Edo	<i>ma</i> (non-past), <i>i</i> (past)	<i>i ghi...</i>	<i>ma he</i>
ST	<i>na...fa</i>	<i>nanta~nantan...fa</i>	<i>naxi...fa</i>
NG	<i>na...wa</i>	<i>na tô...wa</i>	<i>na si...wa</i>
FA	<i>na...f</i>	<i>na tan...f</i>	?
LU	<i>...fa</i>	<i>...ma fa</i>	<i>maxi...fa</i>

The following Edo examples can be contrasted with the examples of preverbal negation markers in Santome (Ch. 4, section 4.3.1)

- (99) *E i ghi yo ugbo ghe hia.* (Edo, Agheyisi 1986: 58)
 He **NEG anymore** go farm time all
 ‘He does not go to the farm at all time anymore.’

- (100) *I ma he kpao.* (Ibidem, p. 59)
 1SG **NEG yet** leave
 ‘I haven’t left yet.’

The morphosyntax of the preverbal negation markers is clearly reminiscent of Edo and not of Bantu or Portuguese. For the same purpose, Kikongo, for example, uses adverbs with greater mobility and which occur mostly in clause initial or post-verbal position (Bentley, 1887). In Portuguese, the adverbs *já* and *ainda* precede the negation marker and *mais* occurs typically in postverbal position, the construction *não mais* ‘no longer’ being stylistically marked.¹²⁹

4.5.3. Western Bantu (area H)¹³⁰

If transfer played a role in the marked negation patterns of the GGC, it will be shown western Bantu languages are a plausible source for the syntax of the final marker (Güldemann & Hagemeyer 2006). Bantu languages are not homogeneous with respect to the syntax of negation. According to Westphal (1958), and simplifying a little bit, basically three patterns are found, namely:

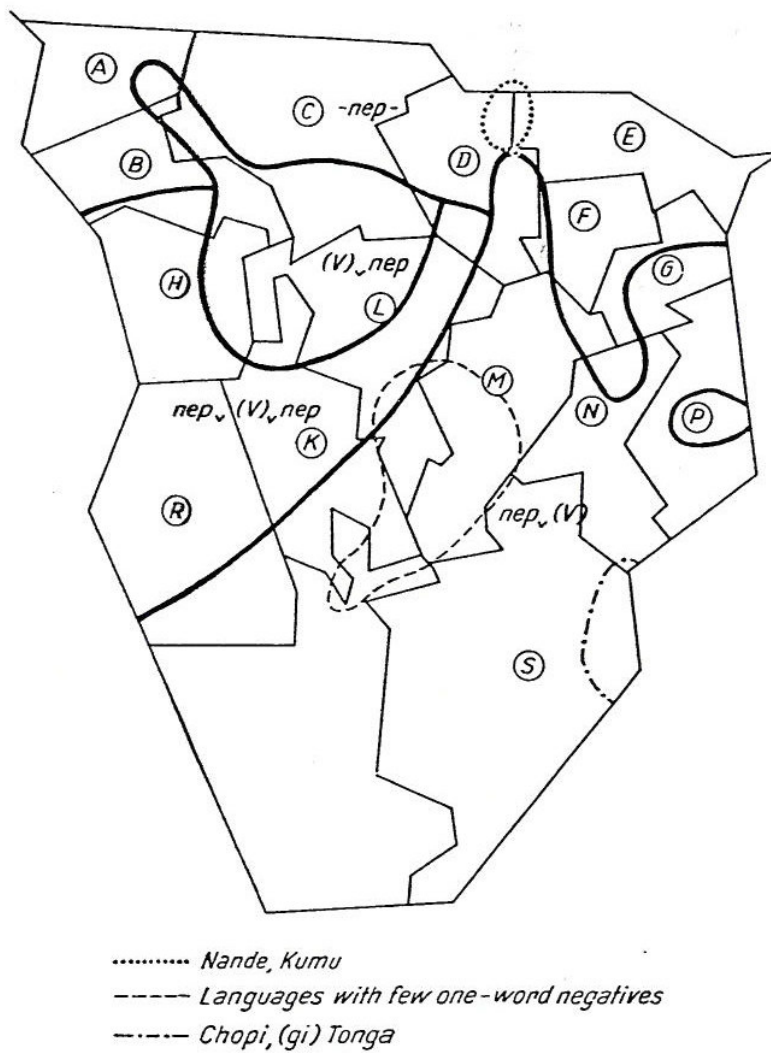
¹²⁹ For historical data, I consulted the data in the electronic database *Corpus Informatizado do Português Medieval* (CIPM) at Universidade Nova de Lisboa.

¹³⁰ This section is part of joint work with Tom Güldemann (Güldemann & Hagemeyer 2006), to whom I am greatly indebted for glossing the Bantu examples.

- i) preverbal negation, expressed by an affix which can be pre-initial or post-initial, depending on whether it follows the subject clitic;¹³¹
- ii) preverbal negation as in (i) with an additional postverbal marker;
- iii) an exclusive final negation marker.

These types have the geographic distribution shown in Map 1.

Map 1: Distribution of types of negation over the Bantu language zones (Westphal 1958)



¹³¹ For a detailed analysis of pre-initial and post-initial negation in Bantu, I refer the reader to Güldemann (1996).

The relevant Bantu area for the formation of the GGC is area H, since it comprises the varieties of Kikongo and Kimbundu. Upon inspection of the map, it follows that the typical pattern in area H consists of a preverbal and postverbal marker. However, there is a significant difference between Kimbundu and Kikongo in this respect, since the former language exhibits a postverbal marker that cliticized on the verbal complex, as illustrated in (101), whereas Kongo languages have a truly final marker, in example (102):

- (101) *Muene kana ka-ri-ê xitu.* (Kimbundu, Chatelain 1888: 147)
 1:DEM NEG 1:PST-eat-NEG meat
 ‘She didn’t eat meat.’
- (102) *Ongue cucuzitissa n Peteleco.* (Kongo; Guinness 1882: 85)
ongwe ku-ku-zitissan Petele ko
 you NEG:2S-FUT-love? Peter NEG
 ‘Thou shalt not love Peter.’

Thus the syntactic position of the Neg2 in Kimbundu superficially resembles the position of French *pas*, even though these elements differ of course with respect to their categorial nature (clitic head vs. specifier). The similarities with negation patterns in Kongo languages are not surprising in the light of other evidence. As Ferraz (1979) shows, the amount of Kongo lexicon in Santome is far more significant than the amount of Kimbundu lexicon. The lexical argument does not hold for Ngola, which retained significant portions of Kimbundu lexicon, but it was shown in section 4.3.3 that this language exhibits essentially the same final negation patterns as in Santome. Under a relexification scenario by Kimbundu, as proposed by Lorenzino (1998), it follows that previously existing syntactic structures did not suffer drastic changes in Ngola when compared to the other GGC, reflecting therefore pre-Kimbundu features (Hagemeijer 1999, 2005b; Hagemeijer & Parkvall 2001). In addition to the typological evidence, genetic evidence corroborates the pre-Kimbundu Bantu layer in the Ngola-speaking population, since its Mt-DNA reveals old lineages that can be traced back to the Kongo region (Rocha et alii under revision).

Being aware of the fact that there is variation within the Kongo language complex, I will provide some data from Bentley (1887), who described the Kikongo variety of São Salvador, which is arguably a relevant slavery area for the island of S.

Tomé. First, this variety exhibits a Neg1...Neg2 pattern, here a negative clause with object, a negative clause with object and adjunct, and a negative imperative clause.

(103) *O mpangi ame ka-sembidi nkanda ko.* (Bentley, 1887: 995)

DEF brother POS NEG:1-buy:PAST book NEG

‘My brother did not buy a book.’

(104) *Kw-endi malembe ko.* (Ibidem, p. 607)

NEG:2SG-go:?IMPER slowly NEG

‘Do not go slowly.’

When Neg1 occurs in the embedding and the embedded clause, Neg2 is not repeated in final position or anywhere else.

São Salvador (H16a)

(105) *Ku-m-pangi diau adimosi ne ki mfumu eno ko.*

NEG:2S-1S-treat:?IMPER like NEG:1S chief POS NEG

‘Do not treat me as if I were not your chief.’ (Bentley 1887: 1032)

Second, Neg2 can be followed by discourse particles, as in Santome. In the following example, *kwe* stands for the contraction of final negation marker *ko*+interrogative particle.

São Salvador (H16a)

(106) *Kw-endi kwe.*

NEG:2S-go NEG-INT

‘Are you not going?’ or ‘Do not go!’ (Bentley 1887: 146)

The negation patterns in the examples above are all identical to what can be observed in Santome. Note, however, that there are also differences. Constituent negation, for examples, requires Neg1 and Neg2 in São Salvador Kikongo, but only Neg2 in Santome.

- (107) *Ke yandi ko, ngâtu mono.* (Bentley 1887: 1018)
 NEG 3SG NEG nor 1SG
 ‘Neither he nor I.’

Kikongo also has cases where Neg2 is not realized, which Bentley labels the “unnatural negative”, describing it as follows:

“When [preverbal] *ke* is used without the second particle of negation [*ko*] the verb appears generally, if not always, in the subjunctive mood. It expresses delight, or surprise, or disappointment, or disapproval, or disgust at the non-fulfilment of what one had feared, expected, hoped for, or considered ought to have been done, or the uselessness of attempting again what has only proved fruitless or useless hitherto”. Bentley (1887: 607)

This distinction is not a generalized property throughout the Kongo cluster however (Kamba Muzenga 1981: 44f). The examples provided by Bentley are not sufficient for a thorough comparison with Santome, but suggest only a partial overlap, for instance in the case of purpose clauses:

San Salvador (H16a)

- (108) *Edika-tung-idi e kozo e ngandu ke zadia e nkombo zandi.*
 ? 1:PST-build-PST fence crocodile.10 NEG 10CONS:eat goat POS
 ‘He built a stock-yard fence, so that the crocodiles should not eat his goats.’
 (Kongo; Bentley 1887: 917)

- (109) *Toma kanga e nkombo zau ke ji-tayi.*
 carefully tie goat.10 10:DEM NEG 10-escape
 ‘Tie the goats carefully, lest they run away.’ (Kongo; Bentley 1887: 608)

Finally, a link between polar item *nê* ‘and, nor, inclusively’, discussed in section 3.1 of this Appendix and Kongo languages is suggestive. Bentley (1887: 968) mentions that “*musungula* means ‘especially, as well as, as well, also’, when it connects with positive sentences, but ‘neither, nor, certainly not’, when it connects with negative clauses.”

- (110) a. *Awonso bekwenda, musungula yandi.* (Kongo; Bentley 1887: 608)
 2:all 2:go CONJ 3SG
 ‘All go, and he will also.’
- b. *Ke ngeye ko, musungula yandi.* (Ibidem)
 NEG 2SG NEG CONJ 3SG
 ‘Not you, and certainly not he.’

All in all, given the typological markedness of discontinuous negation as the default strategy, I assume that the negation patterns observed in Santome and the GGC in general reflect a primary contribution from Kongo languages. In fact, much of the variation found in Kongo languages is also found in the GGC, suggesting that substantial transfer may have taken place with a perhaps more limited role for internal development. The transfer scenario gains strength from the fact that in the few other creole languages with final and/or discontinuous negation this feature is typically correlated with the impact of a given substrate:

- (i) Berbice Dutch – Ijoid (Kouwenberg 1994)
- (ii) Palenquero – Kongo (Schwegler 2006)
- (iii) Afrikaans – Khoekhoe (Den Besten 1986)

Yet, it was argued in section 4.4 that the final marker is an old feature that must have existed in the proto-creole of the Gulf of Guinea prior to diffusion. Therefore it cannot be excluded that *fa* is a very early pre-Kongo element that crystallized with its current syntax due to the impact of the Kongo stratum. However, any of these hypotheses is compatible with the view that Kongo is not a substrate but rather a contact language in the formative process of the GGC whose major contribution was at a phonological level and not at a syntactic level. This means that Kongo languages did not play the same role as Ijoid for Berbice Dutch or Kongo for Palenquero. Thus the discontinuous/final negation pattern in the GGC is arguably the result of contact between an early founder creole, with a crucial role for Portuguese and the Edoid cluster, and Kongo languages.

5. Summary

This Appendix addressed several residual aspects of negation in Santome. In addition to the typologically marked standard negation pattern discussed in Chapter 4, it was shown that Santome is a strong NC language. The arguments presented in section 2 demonstrated that there are no compelling reasons to believe that N-words are not intrinsically negative. A similar treatment of minimizers is warranted. In section 3 it was shown that Santome also exhibits polarity items. This was briefly shown for the case of coordination conjunctions *nê* and correlative *nê...nê* and adverb *antawo*. From section 4 it followed that there must be a diachronic relation between the final negation marker *fa* and emphasis marker *fa~fan*. Moreover, the discontinuous negation patterns can be reconstructed for all the GGC, showing it is an old feature that was present in the proto-GGC. Jespersen's cycle can account for the differences between the GGC with respect to the standard negation patterns. If we compare the GGC creoles in general, and Santome in particular, with the strata that were present during the formative stages of the proto-creole, it becomes clear that its hybrid nature still shows typological vestiges of each layer.

5. CONCLUSIONS

5.1. Synthesis of the findings

For several reasons, stigmatization of various types still plays a role in many creole societies, affecting individuals and speech communities. The situation on S. Tomé is no exception. On many occasions, during my fieldwork, native speakers were ashamed to admit that they were fluent in their native tongue and would claim that Santome is a dialect without rules, not a language. No matter what particular factors belie this type of attitude towards a language, the attitude is part of a negative cycle that needs to be inverted.

The issue of linguistic complexity has accompanied the field of creole studies for the good and the bad ever since the pioneering work by Coelho and Schuchardt. To make a long story short, more than two decades ago Bickerton (1981, 1984) argued that the alleged strong similarities between creole languages are evidence that drastic restructuring, or a genetic break in transmission in the sense of Thomason & Kaufman (1988), uncovers the human language faculty. Despite extensive critique of the LBH, Bickerton's view on creolization has been restored to a significant degree by McWhorter's (1998, 2001, 2005), who proposes a creole prototype that relates directly to the alleged pidgin origins of creole languages. According to this model, creole languages lack certain properties due to being 'young' languages. DeGraff (2001a, 2001b), in particular, has argued against this hypothesis on the grounds that there is nothing about creoles which cannot be explained by language acquisition devices of UG. DeGraff (1999) therefore proposed a Cascade of L1-L2 acquisition, whereby both SLA by adults *and* nativization in the sense of Bickerton play a crucial role.

This dissertation contributes to a better understanding of Santome, and especially its syntax. It was shown that previous accounts of the topics addressed here lacked descriptive and explanatory adequacy. As a result, this creole displays much greater complexity in the relevant domains than could be predicted from the available sources. This complexity readily carries over to other domains of Santome not studied here, such as serialization (Hagemerijer 2000, 2001), and, I am sure, will also be found in yet other unknown or poorly known domains of its grammar. The prototype in the sense of McWhorter, for instance, looks for features for which creole languages are negatively specified. However, not only do these features fail to define creoles as a

linguistic class, but there is also not much point in searching for absent linguistic features when the search for existing features is often still in its early stages. It is also predictable that the increasing knowledge about individual creoles will make the quest for a linguistically identifiable class more and more difficult.

Santome and its sister creoles can be fully explained by drastic restructuring due to a specific type of language contact. Although it is still a relatively young language in the prototypical sense, if it were to still exhibit identical properties within one thousand or two thousand years and if nobody knew its origins, what would it mean, after all, to be a 'young' language? It is beyond any doubt that Santome exhibits a significant amount of substrate features. How else can it be explained that many serializing structures appear to be directly calqued on Edo (Hagemeyer 2005b), sometimes even retaining an Edo verb in the most grammaticalized slot? Furthermore, in this dissertation, it was also shown that the typologically marked negation patterns show strong resemblances with negation in the Kongo cluster. Negative Concord, on the other hand, clearly reveals the presence of Portuguese. The opposition between strong and weak pronouns can be assigned to Bantoid or Edoid. The phonological cliticization processes of object pronouns, especially of 3sg, show strong similarities with the same facts in Edo. Yet, it can easily be demonstrated that Santome is by no means a relexification (Lefebvre 1998) of any particular language. Language contact and internal development come in many forms and degrees and dictate as many different outcomes. In the uniformitarian sense so often invoked by DeGraff, everything is special about Santome, but there is really nothing special about Santome.

In what follows, I will concisely address the specific findings of each of the three chapters.

Chapter 2

The traditional claim made in previous work according to which Santome exhibits weak and strong subject pronouns was essentially correct but clearly in need of refinement. Moreover, it was shown that existing descriptions of the pronominal system were inadequate: some forms had been ignored altogether, whereas others were described incompletely. Above all, however, it was argued that the properties of pronouns in this creole rely heavily on the interplay at the syntax-phonology interface. Linguistic change in the pronominal system, especially with respect to weak 1sg *n*, suggests that this

interface must also include morphology if further weakening of the paradigm takes place and pronouns become reanalyzed as agreement markers.

It followed from their phonological and syntactic properties that pronouns have strongly individualized characteristics on a scale that goes from very weak, i.e. almost syntactic clitics, to very strong pronominals. In fact, weak 1sg *n* is only a small step away from incorporation into V and its extended functional categories, whereas strong 3sg *êlê* is at the opposite end of the scale, radically avoiding weak syntactic environments and, as a consequence, shunning standard Case-marking.

In between the two ends of the scales referred to in the previous paragraph, a significant portion of the pronouns occur in both weak and strong syntactic environments. In order to account for these cases, I considered these pronouns as being syntactically underspecified. The phonological split between clitic and non-clitic pronouns then cuts through the syntactic tripartition into weak, strong and unspecified and yields a finer-grained typology of the pronominal system. Interfacing phonology and syntax has the advantage that one is now able to identify, for instance, strong reduced forms such as *am*, underspecified phonological clitics such as 2sg *ô* and 3pl *nen*, as well as weak non-clitic expletive *kwa*, etc.

For little more than a decade now, an intense debate has been taking place within creole studies about the status of subject pronouns. More specifically, it has been claimed for a number of creole languages that they exhibit syntactic clitics, i.e. heads adjoined to T°. The direct implication of such an analysis is that these agreement markers license a null pronoun in [Spec,TP]. Although it is uncontroversial that rich morphology is not the exclusive licensing factor for *pro*-drop, which is particularly clear in the case of discourse-bound *pro*-drop in Mandarin, creole languages would arguably enrich the typology of *pro*-drop languages. Since most creole languages are strongly isolating languages, the development of agreement markers out of weak pronouns would be one of the first signs of morphological complexification.

In the case of Santome, I provided evidence that subject pronouns cannot be considered syntactic clitics at this stage and therefore Santome does not pattern among the group of creole languages for which *pro*-drop has been invoked. This claim is underscored by the

fact that cases of non-argumental *pro*-drop are extremely limited in this language and the absence of null expletives is a strong indication that a language also lacks *pro*-drop. From this point of view, Santome is thus a typologically well-behaved language.

Chapter 3

The extended VP in most creole languages consists of the predicate and lexicalized preverbal markers that encode TMA. This core feature has played an important role in Bickerton's Language Bioprogram Hypothesis and today, 30 years later, still plays a central role in creole studies. The debate on TMA in Santome goes back as far as the 19th century but generally focuses on the core TMA-markers (*ka*, *ska*, *tava*) and their core properties. In this chapter, I provided new data that adds to the complexity of the Santome TMA-system.

The first part of the chapter examines whether Santome exhibits verb movement. In addition to standard tests, such as adverb placement and quantifier float, I propose two language-internal tests, namely the properties of DOCs and the specific properties of comitatives with respect to goal arguments when they occur with the allomorphs *be* and *ba* 'to go'. Although quantifier float is not a relevant test in Santome because quantifiers do not float, the other constructions can readily be derived with the verb in a static position. This is also consonant with the fact that TMA-markers sit in lexically spelled-out projections and with the conclusion in Chapter 2 that object pronouns are phonological clitics. The isolating character of the language and its rather rigid syntax suggest that the post-syntactic component plays an important role. In addition to the phonology of pronouns, the relation between aspect marker *ka* and the verb and between *sa* and *ka* in the progressive construction shows that the post-syntactic merger is active in different domains.

The second part of the chapter focused on TMA. I argued that Santome is a relative tense language and that stativity is sensitive to a distinction between property predicates, on the one hand, and habitual and stage level predicates, on the other. It was also shown that so-called TMA-markers may differ substantially with respect to their morphosyntactic properties. Aspect marker *ka*, for instance, is a purely functional item that behaves like a bound morpheme on the verb. A number of aspectual constructions that have not been discussed in the previous literature (e.g. *sa ska*, *ka ska*, etc.) strongly

suggest that aspect has been an area of language change. To account for the full range of aspectual data, I proposed that, in addition to AuxP, two AspP projections are required in the structure. The latter projection mediates between TP and AspP or the AspPs in imperfective constructions and is also necessary to distinguish between imperfective constructions, on the one hand, and the perfective of dynamic predicates and the past of stative predicates, on the other. AspP hosts imperfective elements such as *ka*, *ska* or *di* and perfective elements, crucially the \emptyset -marker. The imperfective/perfective distinction correlates with a temporal distinction between present and past, respectively. Tense marker *tava*, which still maintains most of its lexical properties, can be analyzed as a single lexical item with the feature [Past]. This feature amalgamates with the temporal information provided by AspP. For instance, perfective aspect + [Past] triggers a past-before-past reading. When tense is not overtly realized, TP inherits the temporal information from AspP.

The final part of Chapter 3 is dedicated to mood and modality. First, it is shown that Santome has a functionalized mood marker that occurs immediately left of the tense marker. Therefore, Santome, like Ngola and Lung'ie, exhibits an MTA-system rather than a classic TMA-system. It was also shown that it has a functional projection ModP in the low left periphery where at least some modals were shown to occur.

Chapter 4

Santome exhibits a typologically marked negation pattern consisting of a preverbal negation marker *na* and a strongly final marker *fa* which I labeled Neg1 and Neg2. Strongly final means that it occurs to the right of internal arguments, adverbials (XPs) and complement clauses, but also to the right of certain adjunct clauses, such as final relative clauses and circumstantial negative clauses. However, there are also adjunct clauses that block Neg2, such as conditional clauses and reason clauses. Furthermore, I reported a number of mostly irrealis contexts where Neg2 was absent and a number of environments of contrastive negation where Neg1 was lacking. It was shown that the final marker belongs to the same clause as the preverbal marker.

First, I argued that both Neg1 and Neg2 head their own NegP, a proposal that has also been made for a few unrelated languages such as Gengbe and Afrikaans. NegP1, I claimed, is a functional projection that immediately dominates the MTA-system. The

position of NegP2 in discontinuous negation languages is more controversial. Most analyses propose that Neg2 (*fa*) heads a NegP2 or a PolP that sits structurally higher in the clause than NegP1. I departed from this type of analyses because they fail to establish a c-command relation between Neg1 and Neg2. I proposed that NegP2 is a truly negative projection in the lower parts of the TMA-system. Both NegPs were claimed to Agree at distance. I further assumed that a mechanism of AspP-raising to [Spec,NegP2] accounts for the correct surface order of negative sentences. In fact, this proposal is able to subsume a wide array of facts about clause structure in Santome. The abovementioned differential behavior of adjunct clauses with respect to Neg2 reflects the degree of peripherality: low adjuncts adjoin to VP and AspP, whereas high adjuncts adjoin to a high projection, presumably CP. Low adjuncts are therefore in the scope of negation and pied-piped to [Spec,NegP2] as part of the raised AspP. The distribution of negation in coordinate structures also supports the proposed analysis. Neg2 has to be repeated in each conjunct when ConjP establishes a coordination at the level of NegP, but only occurs once in sentence final position when VPs are coordinated. This is fully predicted after raising of the conjoined VPs to [Spec, NegP2].

In the appendix to this chapter, I showed that Santome exhibits both NC and cases of polarity. From a diachronic and comparative perspective, discontinuous negation can be reconstituted in the GGC, which makes this an old feature. At present, it can be seen that each GGC is at a different stage of Jespersen cycle if we assume that they all derived from a single proto-GGC. Finally, I argued that vestiges of the different strata that contributed to the negation patterns during the formative stages of Santome (i.e. the proto-GGC) are still visible in the present day language. NC, including minimizers, must be related to Portuguese, even though Old Portuguese exhibited NPIs where contemporary Santome shows strong NC. I also argued that the tripartite preverbal negation system is reminiscent of Edo, whereas the final marker bears strong resemblances with Kongo languages.

5.2. Issues for further research

The fact that Santome continues to be an understudied language in all domains makes it fertile ground for further research on the topics of this dissertation and beyond. During the preparation of this work, I came across many issues that are worthy of future research:

(i) Adverb *ten~tembeten* ‘also’ and weak pronouns

In Chapter 2 it was shown that even the weakest pronoun of the paradigm and expletives are able to precede this adverb, but not other adverbs with a focalizing function. A careful study of the functions of *ten~tembeten* is required, since it occurs at least as an inclusive adverb and as a Topic marker. There are at least two different hypotheses that should be explored

- (i) the special properties of *ten~tembeten* are related to its unique predicative function;
- (ii) the special properties of *ten~tembeten* are related to underspecification for its X°/XP status.

Hypothesis (i) derives from the fact that this item, different from the other focalizing adverbs, has two scope directions, namely over the preceding constituent or over the predicate. Moreover, in VP-internal position *ten~tembeten* is the only adverb that is able to outscope negation. Therefore it is suggestive that there is a link between its predicative function (Matos 1992) and the licensing of weak forms, even though it was shown that weak forms could also precede *ten~tembeten* when it takes narrow scope over the subject.

Hypothesis (ii) follows Cardinaletti & Starke, who suggest that syntactic deficiency may also reach into domains other than pronouns. This proposal has for instance been made by Castro & Costa (2002) for certain adverbs in Portuguese.

(ii) Aspectual and modal ‘auxiliaries’

Apart from the core TMA-markers, several modal and aspectual items are able to precede the verb, as illustrated in the following examples:

Temporal-aspectual elements

- (1) *Inen **kia** bêndê kani kasô da non.*
3PL almost sell meat dog to 1PL
‘They almost sold dog meat to us.’
- (2) *Ê **bila** xê.*
3SG turn leave
‘He left again.’

- (3) *Ê ka bi môlê.*
 3SG ASP come die
 ‘He will die.’ / ‘He’s going to die.’

Modal elements

- (4) *Bô toka tê ãa ke bô.*
 2SG should have a house POS
 ‘You should have your own house.’
- (5) *Bô pôdja tava fla.*
 2SG should/could TNS speak
 ‘You should/could have spoken.’
- (6) *N tê be.*
 1SG have go
 ‘I have to/must go.’

The full inventory of these items is, of course, more extensive. Preliminary syntactic testing of these items suggests that these items:

- (i) are on a scale that ranges from functional to lexical;
- (ii) exhibit raising properties;
- (iii) often fill in the role of adverbs in European languages.

Property (i) has for instance been addressed by Gonçalves (1996) for European Portuguese. Examples (1), (3) and (6), for instance, respond negatively to most tests that can be used to identify verbs, whereas *bila* in (2) and other items not mentioned here still exhibit a great deal of verbal properties, including the ability to take TMA-markers. *Toka* and *pôdja* in (4) and (5) express necessity and behave in a syntactically similar way. These two items are highly specific in that they can be preceded and followed by tense-aspect material, but not simultaneously. *Ka bi* in (3) is a periphrastic future marker. Alternatively, *ska bi* (progressive+come) and *ka ba* (habitual+go) can also be used for future reference, with fine-grained semantic distinctions. This type of future marking bears a resemblance to the type of constructions studied in Cardinaletti & Giusti (2001), who propose a semi-lexical analysis.

Property (ii) is important in the discussion on raising and control, which has not been studied at all for Santome. Property (iii) might be an interesting test-case for the Cartographic Approach, especially Cinque (1999). One may pose the question of whether some languages exhibit adverbs as specifiers, whereas others lack specifiers and only use the head in the cartography of functional projections. Additionally, it should be investigated whether so-called core TMA-markers can also be integrated in this type of analysis (e.g. Durrleman 2000) and whether there are mixed languages in the sense above.

A detailed study of the properties of these items should shed light on how they project in syntax. Do some of these items head the projections proposed in Chapter 3 for the representation of MTA or is additional structure required?

(iii) Negation

The appendix to Chapter 4 is an immediate reflex of the need for future research into several aspects of negation. One of these aspects would definitely be a comparison between polarity in the four GGC in relation to the standard negation patterns in each creole. It was, for example, shown that N-words in Santome correspond arguably to polarity items in Fa d'Ambô and Ngola. Apart from the brief description in the appendix, polarity has not been explicitly studied for any of the GGC, yet appears to be a promising research domain.

One of the aspects of polarity concerns the relation of Neg2 with strong and weak intensional predicates (SIP / WIP) (e.g. Quer 1998, Borgonovo 2002). I have tested a few cases and the results show interesting similarities with the behaviour of subjunctive/indicative in identical environments in Romance languages. Embedding under an SIP suspends the occurrence of *fa* in Santome, as shown in (7) and (8).

- (7) *pa aman-pasa pa a na fla kuma non ska da mali ngê (*fa).*
 for day after tomorrow for IMP NEG say that 1PL ASP give bad people
 '(...) so tomorrow they won't say that we are speaking badly of people.
- (8) *N (na) mêsê pa bô na kunda kuma ê pô ka fe stluvisu (*fa).*
 1SG NEG want for 2SG NEG believe that 3SG canASP do job
 'I (don't) want you to not think that he can do the job.'

Note that in (7) and (8) Neg1 occurs in a domain that selects a complement clause, which is normally a context that obligatorily triggers Neg2. Emotional factives, on the other hand, are WIPs and show a different effect, irrespective of Neg-raising, in (9) and (10).

(9) *N tê pena ku bô kunda kuma ê na sa klupadu *(fa).*

1SG have pity that 2SG believe that 3SG NEG be guilty

'I regret that you think he is not guilty.'

(10) *N tê pena ku bô na kunda kuma ê sa klupadu *(fa)*

'I regret that you don't think he is guilty.'

Since it was already demonstrated that Neg2 did not occur in subjunctive-like environments, this type of structure constitutes an interesting point of departure toward a better understanding of Neg2 and the relation between Neg1 and Neg2.

References

- Aboh, E. O. (2004). *The morphosyntax of complement-head sequences: clause structure and word order patterns in Kwa*. New York: Oxford University Press.
- _____ (forthc.). "On the right edge c-type negation marker". In E. O. Aboh & J. Essegbey (eds.) *Studies in the Syntax of Kwa languages*.
- Adone, D. (1994). *The acquisition of Mauritian Creole*. Amsterdam: John Benjamins.
- Agheyisi, R. N. (1990). *A grammar of Edo*. Unesco.
- _____ (1991). "The tense systems of Nigerian languages and English". In O.E. Essien (ed.) *Afrikanistische Arbeitspapiere*, 27: 11-42.
- Alcântara, C. (Ms.). *A bamu fla Santome*. S. Tomé.
- Alexandre, N. & T. Hagemeyer (2002). "Pronomes resumptivos e abandono de preposição nos crioulos atlânticos de base lexical portuguesa". In *Actas do XVII Congresso da Associação Portuguesa de Linguística*, 17-30.
- _____ (in press). "Bare nouns and the nominal domain in Santome". In M. Baptista & J. Guéron (eds.) *Noun phrases in creole languages: a multi-faceted approach*. Amsterdam: John Benjamins.
- Alexiadou, A. (1994). *Issues in the syntax of adverbs*. PhD Dissertation, University of Potsdam.
- Andersen, R. W. & Y. Shirai (1996). "Primacy of aspect in first and second language acquisition: the pidgin/creole connection". In W. C. Ritchie & T. K. Bhatia (eds.) *Handbook of second language acquisition*. San Diego, CA: Academic Press, 527-570
- Auger, J. (1994). *Pronominal clitics in Québec colloquial French: a morphological analysis*. PhD Dissertation, University of Pennsylvania.
- Baker, M. (1988). *Incorporation*. Chicago: University of Chicago Press.
- _____ (1996). "On the structural positions on themes and goals". In J. Rooryck & L. Zaring (eds.) *Phrase Structure and the Lexicon*. Dordrecht: Kluwer, 7-34.
- _____ (1997). "Thematic roles and syntactic structure". In L. Haegeman (ed.) *Elements of Grammar*. Dordrecht: Kluwer, 73-137.
- Baptista, M. (2002). *The syntax of Cape Verdean creole: the sotavento varieties*. Amsterdam: John Benjamins.
- Barbiers, S. (2002). "Microvariation in negation in varieties of Dutch". In S. Barbiers, L. Cornips & S. Van der Kleij (eds.) *Syntactic microvariation*. Meertens Institute Electronic Publications in Linguistics.
- Barrena, R. (1957). *Gramatica annobonesa*. Madrid: Consejo de Investigaciones Cientificas.
- Bayer, J. (1996). *Directionality and logical form: on the scope of focussing particles and Wh in situ*. Dordrecht: Kluwer.
- Becker, A. & T. Veenstra (2003). "Creole prototypes as basic varieties and inflectional morphology". In C. Dimroth & M. Starren (eds.) *Information structure and the dynamics of language acquisition*. Amsterdam/Philadelphia: John Benjamins, 235-264.
- Bell, A. J. (2004). *Bipartite negation and the fine structure of the negative phrase*. PhD Dissertation, Cornell University.
- Bentley, W. H. (1887). *Dictionary and grammar of the Kongo language*. London: Baptist Missionary Society.
- Bickerton, D. (1975). *Dynamics of a creole system*. Cambridge: Cambridge University Press.
- _____ (1981). *Roots of language*. Ann Arbor: Karoma.

- _____ (1984). "The language bioprogram hypothesis". In *The behavioral and brain sciences* 7: 173-221.
- Bobaljik, J. D. (1995). *Morphosyntax: the syntax of verbal inflection*. PhD Dissertation, MIT, Cambridge (MA).
- _____ (2000). "The rich agreement hypothesis in review". Ms., McGill University.
- Bobaljik, J. D. & H. Thráinsson (1998). "Two heads aren't always better than one". In *Syntax* 1: 37-71.
- Bonfim, F. *A Liberdade* II (25, 26), III (28).
- _____ [unpublished pamphlets, 1940s/1950s]
- Borer, H. (1986). "I-subjects". In *Linguistic Inquiry* 17: 375-416.
- _____ (1989). "Anaphoric agr.". In O. Jaeggli and K. J. Safir (eds.) *The null subject parameter*. Dordrecht: Kluwer Academic Publishers, 69 - 109.
- _____ (1997). "Functional projections: at the interface of acquisition, morphology and syntax". In G. Matos et alii (eds.) *Interfaces in linguistic theory*. Lisboa: Edições Colibri.
- Bowers, J. (1993). "The syntax of predication". In *Linguistic Inquiry* 24, 591-656.
- Borgonovo, C. (2002). "Mood and focus". In *Cuadernos de Lingüística IX*. Instituto Universitario Ortega y Gasset: 43-63.
- Bosque, I. (1990). *Indicativo y Subjuntivo*. Madrid: Taurus.
- Brandi, L. & P. Cordin (1989). "Two Italian dialects and the null-subject parameter". In O. Jaeggli & K. Safir (eds.) *The Null-Subject Parameter*. Dordrecht: Kluwer, 111-142.
- Bresnan, J & S. Mchombo (1987). "Topic, pronoun and agreement in Chichewa". In *Language* 63(4), 741-82.
- Bruyn, A. & P. Muysken & M. Verrips (1999). "Double-object constructions in the creole languages: development and acquisition". In DeGraff (ed.) *Language change and creation: creolization, diachrony, and development*, 329-374.
- Büring, D & K. Hartmann. (2003). "The syntax and semantics of focus-sensitive particles in German". *Natural Language and Linguistic Theory* 19: 229-281.
- Byrne, F. (1985). "Pro_{prox} in Saramaccan". In *Linguistic Inquiry*, 16.
- Cadely, J. R. (1994). *Aspect de la phonologie du Créole Haïtien*. PhD Dissertation, Université du Québec à Montréal.
- Cardinaletti, A. & G. Giusti (2001). "'Semi-lexical' motion verbs in Romance and Germanic". In N. Corver & H. van Riemsdijk (eds.) *Semi-lexical categories*. Berlin: Mouton De Gruyter, 371-414.
- Cardinaletti, A & M. Starke (1994). *The Typology of Structural Deficiency*. Ms., Universities of Venice and Geneva.
- _____ (1995). "Deficient pronouns: A view from Germanic: a study in the unified description of Germanic and Romance". Ms., Universities of Venice and Geneva.
- _____ (1999). "The typology of structural deficiency: a case study of the three classes of pronouns". In H. van Riemsdijk (ed.) *Clitics in the Languages of Europe* Berlin/New York: Mouton de Gruyter.
- Castro, A. & J. Costa (2002). "Possessivos e advérbios: formas fracas como X^o". In *Actas do XVII Encontro da Associação Portuguesa de Linguística*. Lisboa: APL, 101-111.
- Chatelain, H. (1888-1889). *Grammatica elementar do Kimbundu ou língua de Angola*. Genebra: Typ. Charles Schuchardt.
- Carlson, G. (1977). *Reference to Kinds in English*. PhD Dissertation, University of Massachusetts, Amherst (MA).
- Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

- _____ (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- _____ (1986). *Barriers*. Cambridge/London: MIT Press.
- _____ (1995). *The Minimalist Program*. Cambridge (MA): MIT Press.
- _____ (1999). "Derivation by phase". *MIT Occasional Papers in Linguistics* 18.
- _____ (2001). "Beyond Explanatory Adequacy". *MIT Occasional Papers in Linguistics* 20.
- Cinque, G. (1999). *Adverbs and Functional Heads*. New York: Oxford University Press.
- Coelho, F. A. (1880-6). "Os dialectos românicos ou neolatinos na África, Ásia e América". In *Boletim da Sociedade de Geografia de Lisboa*. I. 2 series 3 (1980), 129-196; II. 3, series 8 (1882), 451-478; III. 6 series 12 (1886), 705-755.
- Colaço, M. (2005). *Configurações de coordenação aditiva: tipologia, concordância e extracção*. Doctoral Dissertation, Universidade de Lisboa.
- Collins (1997). *Local Economy*. In Cambridge/London: MIT Press.
- Comrie, B. (1976). *Aspect: an introduction to verbal aspect and related problems*. Cambridge: Cambridge University Press.
- Comrie, B. (1985). *Tense*. Cambridge: Cambridge University Press.
- Corver, N. & D. Delfitto (1993). "Feature Asymmetry and the nature of Pronoun Movement". Ms., University of Utrecht.
- Costa, J. (1998). *Word order variation: A constrained approach*. The Hague: Holland Academic Graphics.
- _____ (2004). "A multifactorial approach to adverb placement: assumptions, facts, and problems". *Lingua*, 114:6, 711-53.
- Costa, J. & F. Pratas (2004). "Capeverdean creole: some parametric values". In M. Fernández, M. Fernández-Ferreiro, N. Vázquez Veiga (eds.) *Los Criollos de Base Ibérica: ACBLPE 2003*. Madrid: Vervuert/Iberoamericana, 126-136.
- Cunha, L. F. (1998). *As Construções com progressivo no português: uma abordagem semântica*. MA Thesis, University of Porto.
- Curtin, P. D. (1969). *The Atlantic slave trade: a census*. Madison (WI): University of Wisconsin Press.
- Daio, O. (2002). *Semplu: colectânea de provérbios e máximas santomenses*. S. Tomé: Edições Gesmédia.
- De Cat, C. (2005). "French subject clitics are not agreement morphemes". *Lingua*, 108, 1195-1219.
- De Crousaz, I. & U. Shlonsky (2000). "The distribution of a subject clitic pronoun in a Franco-Provençal dialect". Ms., University of Geneva.
- Degraff, M. (1993). "Is Haitian creole a pro-drop language?". In F. Byrne & J. Holm (eds.) *Atlantic meets Pacific: a global view of pidginization and creolization*. Amsterdam: John Benjamins, 71-90.
- _____ (1999). "Creolization, language change, and language acquisition: an epilogue". In DeGraff, M. (ed.) *Language change and creation: creolization, diachrony, and development*. Cambridge: MIT Press, 473-543.
- _____ (2001a). "Against creole exceptionalism". *Language* 79: 391-410.
- _____ (2001b). "On the origin of creole. A cartesian critique of neo-darwinian linguistics." *Linguistic Typology* 5: 213-310.
- _____ (1997). "Verb syntax in, and beyond, creolization". In L. Haegeman (ed.) *The new comparative syntax*. London: Longman, 64-94.
- Demuth, K. & J. Gruber (1995). "Constraining XP sequences". In V. Manfredi & K. Reynolds (eds.) *Niger-Congo Syntax and Semantics* 6. Boston, Boston University African Studies Center.

- Den Besten, H. (1986). "Double negation and the genesis of Afrikaans". In P. Muysken & N. Smith (eds.) *Substrata versus universals in creole languages*. Amsterdam: John Benjamins, 185-230.
- Déprez, V. (1994). "Haitian creole: a pro-drop language?" *Journal of Pidgin and Creole Languages* 9, 1-24.
- _____ (1999). "The roots of negative concord in French and French based creoles". In DeGraff, M. (ed.) *Language creation and language change: creole, diachrony and development*. Cambridge (MA): MIT Press, 375-428.
- Dieck, M. (2000). *La negación en palenquero: análisis sincrónico, estudio comparativo y consecuencias teóricas*. Madrid/Frankfurt: Iberoamericana & Vervuert.
- Diesing, M. (1992). *Indefinites*. Cambridge (MA): MIT Press
- Dietrich, R., W. Klein & C. Noyau (1995). *The acquisition of temporality in a second language*. Amsterdam/Philadelphia: John Benjamins.
- Donaldson, B. C. (1993). *A Grammar of Afrikaans*. Berlin & New York: Mouton de Gruyter
- Dowty, D. (1979). *Word meaning and Montague Grammar*. Dordrecht: Reidel.
- Duarte, I. & G. Matos (2000). "Romance clitics and the minimalist program". In J. Costa (ed.) *Portuguese syntax: new comparative studies*. Oxford: Oxford University Press, 116-142.
- Durleman, S. (2000): "The architecture of the clause in Jamaican creole". *Generative Grammar in Geneva* 1, 189-240.
- Ernst, Th. (2002). *The syntax of adjuncts*. Cambridge: Cambridge University Press.
- Ferraz, L. (1979). *The Creole of São Tomé*. Johannesburg: Witwatersrand University Press.
- Ferraz, L. & M. F. Valkhoff (1975). "A comparative study of São-tomense and Caboverdiano Creole". In M. F. Valkhoff (ed.) *Miscelânea luso-africana*. Lisboa: Junta de Investigações Científicas do Ultramar, 15-39.
- Fiéis, M. A. (2003). *Ordem de palavras, transitividade e inacusatividade. Reflexão teórica e análise do português dos séculos XIII a XVI*. Doctoral Dissertation, Universidade Nova de Lisboa.
- Fuß, Eric (Ms.). "On the historical origin of pro-drop".
- Giannakidou, A. (1998). *Polarity sensitivity as (non)veridical dependency*. Amsterdam: John Benjamins.
- Givón, T. (1982). "Tense-aspect-modality: the creole proto-type and beyond". In P. Hopper (ed.) *Tense-aspect: between semantics & pragmatics*. Amsterdam: John Benjamins, 115-63.
- Gonçalves, A. (1996). "Aspectos da sintaxe dos verbos auxiliares do Português Europeu". In *Quatro estudos em sintaxe do português*. Lisboa: Colibri, 7-50.
- _____ (1999). *Predicados Complexos Verbais em Contextos de Infinitivo não Preposicionado do Português Europeu*. Doctoral Dissertation, Universidade de Lisboa.
- Grimshaw, J. (1991). *Extended Projections*. Ms., Brandeis University.
- Guinness, H. G. (1882). *Grammar of the Congo language as spoken two hundred years ago, translated from the Latin of Brusciotto*. London: Hodder & Stoughton.
- Güldemann, T. (1996). *Verbmorphologie und Nebenprädikationen im Bantu: eine Studie zur funktionalmotivierten Genese eines konjugationalen Subsystems*. Bochum: Brockmeyer.
- Güldemann, T. & T. Hagemeyer (2006). "Negation in the Gulf of Guinea creoles: typological and historical perspectives". In Talk delivered at the ACBLPE meeting. Coimbra.
- Günther, W. (1973). *Das Portugiesische Kreolisch der Ilha do Príncipe*. Marburg an der Lahn: Marburger Studien zur Afrika und Asienkunde.
- Haegeman, L. (1995). *The Syntax of Negation*. Cambridge: Cambridge University Press.

- _____ (2002a). *Anchoring to speaker, adverbial clauses and the structure of CP*. Lille: Université Charles de Gaulle-Lille III Ms.
- _____ (2002b). "Approaches to OV. Derivation of OV Orders and West Flemish Negation". *Progress in Grammar*, Meertens Institute.
- Haegeman, L. & R. Zanuttini (1991). "Negative heads and the neg-criterion". *The Linguistic Review* 8, 233-251
- Hagemeijer, T. (1999). "As ilhas de Babel: a crioulação no Golfo da Guiné". *Revista Camões* 6, 74-88.
- _____ (2000). *Serial verb constructions in São-Tomense*. MA Thesis, Universidade de Lisboa.
- _____ (2001). "Semi-lexicality and underspecification in serial verb constructions". In N. Corver & H. van Riemsdijk (eds.) *Semi-lexical heads*. New York: Mouton de Gruyter, 415-451.
- _____ (2003). "A negação nos crioulos do Golfo da Guiné: aspectos sincrónicos e diacrónicos". *Revista internacional de lingüística iberoamericana*, 2, 161-188.
- _____ (2005a). "Going in the clause: *ba* and *be* in Santome". In N. Smith (ed.) *Journal of Portuguese Linguistics*, 3(2), 71-95.
- _____ (2005b). "The origins of serialization in the Gulf of Guinea creoles". Talk delivered at Creole language structure: between substrates and superstrates, Max Planck Institute, Leipzig.
- _____ (forthc.). "Aspects of discontinuous negation in Santome". In N. Cyffer & E. Ebermann (eds.) *Negation in West-African languages*. Amsterdam/Philadelphia: John Benjamins.
- Hagemeijer, T. & M. Parkvall (2001). "The development of the Gulf of Guinea creoles". Talk delivered at the ACBLPE meeting, Coimbra, Portugal.
- Halle, M. & A. Marantz. (1993). "Distributed morphology". In J. Keyser (ed.) *The View from Building 20*. Cambridge (MA): MIT Press.
- Hoeksema, J. (2000). "Negative polarity items: triggering, scope, and c-command". In R. H. Laurence & Y. Kato (eds.) *Negation and polarity: syntactic and semantic perspectives*. Oxford/Stanford: Oxford University Press, 115-146.
- Hoekstra, T. & R. Mulder (1990). "Unergatives as copular verbs: locational and existential predication". *The Linguistic Review* 7, 1-79.
- Holmberg, A. (1986). *Word order and syntactic features in the Scandinavian languages and English*. PhD dissertation, University of Stockholm.
- Holmberg, A. & C. Platzack (1995). *The role of inflection in Scandinavian syntax*. Oxford: Oxford University Press.
- Holm, J. (1989). *Pidgins and creoles* (vols.1&2). Cambridge: Cambridge University Press.
- Horn, L. R. (1989). *A natural history of negation*. Chicago: The University of Chicago Press.
- Huang, J. (1984). "On the distribution and reference of empty pronouns". *Linguistic Inquiry* 15, 531-574
- _____ (1989). "Pro-drop in Chinese: a generalized control theory". In O. Jaeggli & K. Safir (eds.) *The Null Subject Parameter*. Dordrecht: Kluwer, 185- 214.
- Hyams, N. (1994). "V2 null arguments and COMP projections". In T. Hoekstra & B. Schwartz (eds.) *Language Acquisition Studies in Generative Grammar*. Amsterdam: John Benjamins, 21-55.
- Ionin, T. & O. Matushansky (2002). "DPs with a twist: A unified analysis of Russian comitatives". *Proceedings of FASL 11* (Amherst, MA).
- Jaeggli, O. & K. Safir (1989). "The null subject parameter and parametric binding". In O. Jaeggli & K. Safir (eds.) *The Null Subject Parameter*. Dordrecht: Kluwer, 1-44.

- Johnston, M. (1993). *Because clauses and negative polarity licensing. Proceedings of ESCOL 93*. Ithaca: Cornell University Press.
- Kahrel, P. (1996). *Aspects of negation*. Doctoral Dissertation. Amsterdam: University of Amsterdam.
- Kamba Muzenga, J. G. (1981). "Les formes verbales négatives dans les langues bantoues". In *Annalen menselijke wetenschappen*. Koninklijk Museum voor Midden-Afrika 106.
- Kamp, H. & U. Reyle (1993). *From Discourse to Logic*. Dordrecht: Kluwer Academic Publishers.
- Kayne, R. (1975). *French syntax: the transformational cycle*. Cambridge (MA): The MIT Press.
- _____ (1984). *Connectedness and binary branching*. Foris: Dordrecht, 103-23.
- _____ (1994). *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Koenenman, O. (2000). *The Flexible Nature of Verb Movement*. PhD Dissertation, Utrecht University.
- König, E. (1991). *The meaning of focus particles. A comparative perspective*. London, Routledge.
- Kouwenberg, S. (1994). *A grammar of Berbice Dutch Creole*. Berlin/New York: Mouton de Gruyter.
- _____ (2006). "The status of weak pronouns in Papiamentu". Talk delivered at the ZAS, Berlin.
- Kratzer, A. (1995). "Stage-level and individual-level predicates". In Carlson, G. e F. Pelletier (eds.) *The generic book*. Chicago: University of Chicago Press, 125-175.
- Krifka (1999). "Additive particles under stress". In *Proceedings of SALT*. Cornell: CLC Publications 8, 111-128,
- Ladusaw, W. (1979) "Polarity sensitivity as inherent scope relations," PhD Dissertation, University of Texas, Austin.
- _____ (1992). "Expressing Negation". *Semantics and Linguistic Theory (SALT) 2*: 220-229.
- Laka, I. (1994). *On the syntax of negation*. New York/London: Garland Publishing.
- Laenzlinger, C. (1998). *Adverbs, pronouns, and clause structure in Romance and Germanic*. Amsterdam/Philadelphia: John Benjamins.
- Larson, R. (1988). "On the Double Object Construction". *Linguistic Inquiry* 19, 335-391.
- Lefebvre, C. (1998). *Creole genesis and the acquisition of grammar*. Cambridge (UK): Cambridge University Press.
- Lefebvre, C. and A.-M. Brousseau (2002) *A grammar of Fongbe*. Mouton de Gruyter, Berlin.
- Lehmann, C. & S. Yong-Min (2005), "The functional domain of concomitance". In Lehmann, C. (ed.) *Typological studies in participation*. Berlin: Akademie (Studia Typologica) 6, 9-104.
- Lobo, M. (2002). "On the structural position of non-peripheral adjunct clauses". *Journal of Portuguese linguistics* 1: 83-118.
- _____ (2003). *Aspectos da sintaxe das orações subordinadas adverbiais do português*. Doctoral Dissertation, Universidade Nova de Lisboa.
- Lorenzino, G. (1996). "Afro-Portuguese creole *a*: Its Kwa origins and discourse pragmatics". *African Journal of Languages and Linguistics* 1, 45-67.
- _____ (1998). *Angolar creole Portuguese*. Newcastle: Lincom Europa.
- MacDonald, J. E. (2006). *The syntax of inner aspect*. PhD Dissertation, Stony Brook University.
- Maling, J. (1976). "Notes on quantifier-postposing". *Linguistic Inquiry* 7, 708-718.
- Manninen, S. (2002). "A minimalist analysis of stage level and individual level predicates". *Working Papers in Linguistics, Vol. 1*, The Department of English in Lund.
- Marantz, A. (1988). "Clitics, morphological merger, and the mapping to phonological structure". In M. Hammond & M. Noonan (eds.) *Theoretical Morphology*. San Diego: Academic Press, 253-270.

- _____ (1993). "Implications of asymmetries in double object constructions". In S. Mchombo (ed.) *Theoretical aspects of Bantu grammar*. Stanford: CSLI Publications, 113-150.
- Martins A. M. (1997). "Aspectos da negação na história das línguas românicas (da natureza de palavras como *nenhum, nada, ninguém*)". In *Actas do XII encontro nacional da associação portuguesa de linguística*. Lisboa: APL, 170-210.
- _____ (2000). "Polarity items in Romance: underspecification and lexical change". S. Pintzuk, G. Tsoulas & A. Warner (eds.) *Diachronic syntax: models and mechanisms*. Oxford/New York: Oxford University Press, 232-248.
- Matos, G. (1992). *Construções de eclipse do predicado em português – SV nulo e despojamento*. Doctoral Dissertation, Universidade de Lisboa.
- _____ (2003). "Construções elípticas". In M. H. Mateus et alii (eds.) *Gramática da língua da língua portuguesa*. Lisboa: Caminho, 869-913.
- Matsuo, A. (2001). Asp-shell structure in VP-ellipsis and ACD. In *WCCFL XX*, 386-399.
- Maurer, P. (1992). "L'apport lexical bantou en Angolar". *Afrikanische Arbeitspapiere* 29, 163-174.
- _____ (1995). *L'Angolar: Un créole afro-portugaise parlé à São Tomé*. Hamburg: Helmut Buske Verlag.
- _____ (1997). "Tense-aspect-mood in Principense". In A. Spears & D. Winford, (eds.) *The structure and status of pidgins and creoles*. Amsterdam: John Benjamins.
- _____ (forthc). *Lung'ie. The afro-Portuguese creole of Príncipe island (Gulf of Guinea). Grammar – texts – vocabulary*. London: Battlebridge.
- McCloskey, J. & K. Hale (1984). "On the syntax of person-number inflection in modern Irish". *Natural Language and Linguistic Theory* 1(4): 487-534.
- McWhorter, J. (1998). "Identifying the creole prototype: vindicating a typological class". *Language* 74(4): 788-818.
- _____ (2001). "The simplest grammars are creole grammars". *Linguistic Typology* 5-2, 125-66.
- _____ (2005). *Defining creole*. Oxford: Oxford University Press.
- Melzian, H. J. (1942): *Vergleichende Charakteristik des Verbums im Bini*. Leipzig: Otto Harrassowitz.
- Merchant (2001). "Why no(t)?" Ms., University of Chicago.
- Michaelis, L. (1993). *Toward a grammar of aspect*. PhD Dissertation, University of California, Berkeley.
- Milsark, G. (1974). Existential sentences in English. PhD Dissertation, MIT.
- Moens, M. (1987). *Tense, aspect and temporal reference*. PhD Dissertation, University of Edinburgh.
- Molnarfi, L. (2002). "Die negationsklammer im Afrikaans: Mehrfachnegation aus formaler und funktionaler Sicht". In Abraham, W. (Ed.) *Issues in formal german(ic) typology*. Philadelphia: John Benjamins, 241-279.
- Muysken, P. (1981). "Creole tense/mood/aspect systems: the unmarked case?". In P. Muysken, (ed.) *Generative Studies on Creole Languages*. Dordrecht: Foris, 181-199.
- Negreiros, A. (1895). *Historia ethnographica da ilha de S. Tomé*. Lisboa [no ed.].
- Omoruyi, T. O. (1989). "Negation in Edo". *Afrikanistische Arbeitspapiere* 19, 127-146.
- Oosthuizen, J. (1998). "The final *nie* in Afrikaans negative sentences". In *Stellenbosch Papers in Linguistics* 31, 61-93.
- Ouhalla, J. (1990). "Sentential negation, relativised minimality, and the aspectual status of auxiliaries". *The Linguistic Review*, 7: 83-231.
- Palmer, F. R. (1986) *Mood and modality*. Cambridge: Cambridge University Press.
- Pesetsky, D. (1995). *Zero Syntax*. Cambridge: MIT Press.

- Poletto, C. (2000). *The higher functional field*. Oxford: Oxford University Press.
- Pollock, J. (1989). "Verb movement, UG and the structure of IP". *Linguistic Inquiry* 20, 365-424.
- Posner, R. (1984). "Double negatives, negative polarity and negative incorporation in Romance: a historical and comparative view". In *Annals of the Philological Society*. Oxford: Blackwell, 1-26.
- Post, M. (1995). "Aspect marking in Fa d'Ambu". In Baker Ph. (ed.) *From Contact to Creole and Beyond*. London: University of Westminster Press
- _____ (1997). "Negation in Fa d'Ambô". In R. Degenhardt, T. Stolz, & H. Ulferts (eds.) *Afrolusitanistik – eine vergessene Disziplin in Deutschland?* Bremen: Universität Bremen, 292-316.
- _____ (2000) "Generic and impersonal pronouns in Fa d'Ambô". In E. d'Andrade, D. Pereira, M. A. Mota (eds.) *Crioulos de base portuguesa*. Braga: APL
- Pratas, F. (2002). *O Sistema Pronominal do Caboverdiano (variante de Santiago): questões de gramática*. MA thesis, Lisbon, Universidade Nova de Lisboa.
- _____ (2004) "TP in Capeverdean: an almighty projection?". Talk delivered at the Lisbon workshop on alternative views on the functional domain, Lisbon, Universidade Nova de Lisboa.
- Progovac, L. (1994). *Negative and positive polarity: a binding approach*. Cambridge: Cambridge University Press.
- Quer, J. (1998). *Mood at the Interface*. PhD Dissertation, University of Utrecht.
- Quintas da Graça, A. (1989). *Paga Ngunu*. Empresa de Artes Gráficas, S. Tomé.
- Reinhart T. (1995). *Interface strategies*. Utrecht: Research Institute for language and Speech (OTS).
- Reichenbach, H. (1947). *Elements of Symbolic Logic*. MacMillan, London.
- Revista Cultural* 1, S. Tomé, 1990.
- Rivero, M. L. (1997). "On two locations for complement clitic pronouns: Serbo-Croatian, Bulgarian and Old Spanish. In A. van Kemenade & N. Vincent (eds.) *Parameters of morphosyntactic change*. Cambridge: Cambridge University Press, 170-206.
- Rizzi, L. (1986). "Null objects and the theory of pro". *Linguistic Inquiry* 17: 501-557.
- _____ (1997). "The fine structure of the left periphery". In L. Haegeman (ed.) *Elements of grammar*. Dordrecht: Kluwer, 281-337.
- Robbers, K. (1993). "Some Remarks on Focus and Negation in Afrikaans". Ms., University of Amsterdam.
- _____ (1997). *Non-finite verbal complements in Afrikaans: A comparative approach*. The Hague: Holland Academic Graphics.
- Roberts, I. (1999). "Verb movement and markedness". In M. DeGraff (ed.) *Language Creation and Language Change*. Cambridge (MA): MIT Press.
- Rocha, J. et alii (under revision). "Human microevolution and the Atlantic slave trade: the case study of São Tomé (Gulf of Guinea)". *Current Anthropology*.
- Rohrbacher, B. (1999). *Morphology-driven syntax: a theory of V to I raising and pro-drop*. Amsterdam: John Benjamins.
- Rooryck, J. (2001). "Evidentiality, Part I." *Glott International* 5 (4), 125-133.
- Ross, J. (1967). *Constraints on variables in syntax*. Indiana University Linguistics Club.
- _____ (1969). "Auxiliaries as main verbs". In E. Wood (ed.) *Studies in Philosophical Linguistics Series I*. Evanston: Great Expectations Press, 77-102
- Rottet, K. (1992). "Functional categories and verb movement in Louisiana creole". *Probus* 2, 261-289.

- Sasse, H. J. (2002). "Recent activity in the theory of aspect: accomplishments, achievements, or just non-progressive state?". *Linguistic Typology* 6, 199–271.
- Schang, E. (2000). *L'émergence des créoles portugais du golfe de Guinée*. Doctoral Dissertation, l'Université Nancy 2: Presses Universitaires du Septentrion.
- Schuchardt (1882). "Kreolische Studien I. Über das Negerportugiesische von S. Thomé". *Sitzungsberichte d. Wien. Ak.* 101, II: 889-917.
- _____ (1889). "Beiträge zur Kenntnis des kreolischen Romanisch: IV. Zum Negerportugiesischen der Ilha do Principe". *Zeitschrift für rom. Philologie* 13: 463-475.
- Schwegler, A. (1990). *Analyticity and syntheticity: A diachronic perspective with special reference to Romance languages*. Berlin: Mouton de Gruyter.
- _____ (1991). "Negation in Palenquero. Synchrony". *Journal of pidgin and creole languages* 6, 165-214.
- _____ (2002). "On the (African) origins of Palenquero subject pronouns." *Diachronica* 19, 273-332.
- _____ (2006). "Bantu elements in Palenque (Colombia): anthropological, archeological and linguistic evidence". In *African Re-Genesis: Confronting Social Issues in the Diaspora*, Havisser, J.B. & K. C. MacDonald. London: University College London Press 204-222.
- Schütze, C. (1995). "PP attachment and argumenthood". In *Papers on language processing and acquisition* 26, 95-151.
- Speas, M. (2004). "Logophoricity and the syntactic representation of pragmatic features". *Lingua* 114(3): 255–276.
- Spencer & Zaretskaya (2003). "Stative nominalizations: stative predicates in Russian and their nominalizations". *ERRL* 22.
- Sportiche, D. (1988). "A theory of floating quantifiers and its corollaries for constituent structure". *Linguistic Inquiry* 19: 425-49.
- Sportiche, D. (1996). "Clitic Constructions". In J. Rooryck & L. Zaring (eds.) *Phrase structure and the lexicon*. Dordrecht: Kluwer, 213-276
- Stowell, T. (1993). "The syntax of tense. Ms., UCLA.
- Syea, A. (1993). "Null subject in Mauritian creole and the pro-drop parameter". In F. Byrne & J. Holm (eds.) *Atlantic meets Pacific: a global view of pidginization and creolization*, 91-102.
- Thomason, S. & T. Kaufman (1988). *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Thompson, E. (1996). *The Syntax of Tense*. Ph.D. Dissertation, University of Maryland at College Park.
- Thráinsson, H. (1996). "On the (non-)universality of functional categories". In W. Abraham et alii (eds.) *Minimal ideas. Syntactic studies in the minimalist framework*. Amsterdam/Philadelphia: John Benjamins, 253-281.
- Tomás, G. et alii (2002). "The peopling of São Tomé (Gulf of Guinea): origins of slave settlers and admixture with the Portuguese". *Human Biology* 74, 397-411.
- Travis L. (1984). *Parameters and effects of word order variation*. PhD Dissertation, MIT.
- Valkhoff, M. (1966). *Studies in Portuguese and Creole*. Johannesburg: Witwatersrand University Press.
- Vallduví, E. (1994). "Polarity items, n-words and minimizers in Catalan and Spanish". *Probus* 6: 263-294.
- Veenstra, T. (1996). *Serial verbs in Saramaccan*. The Hague: HIL.

- _____ (2004). "Unaccusativity in Saramaccan: the syntax of resultatives". In A. Alexiadou, E. Anagnostopoulou, & M. Everaert (eds.) *The Unaccusativity Puzzle*. Oxford: Oxford University Press, 269-287.
- _____ (2006). "On typological correlates and pure syntax: evidence from expletives in Romance-based creoles". In H. M. Gärtner et alii (eds.) *Between 40 and 60 for Krifka* (<http://www.zas.gwz-berlin.de/40-60-puzzles-for-krifka/>)
- _____ (to appear). "Syntax pur: expletiva in Papiamentu". In G. Mensching & E. M. Remberger (eds.) *Minimalistische Studien der Romanischen Syntax*. Tübingen: Günther Narr.
- Vikner, S. (1990). *Verb movement and the licensing of NP-positions in the Germanic Languages*. PhD Dissertation, University of Geneva.
- _____ (1995). *Verb movement and expletive subjects in the Germanic languages*. Oxford: Oxford University Press.
- Vila, I. (1891). *Elementos de la gramática ambú ó de Annobón*. Madrid: Imprenta de A. Pérez Dubrull.
- Westphal, E. O. J. (1958). "An introductory comparative study of negation in Bantu." *Mitteilungen des Instituts für Orientforschung* 6, 2: 284-320.
- Zanuttini, R. (1991). *Syntactic properties of sentential negation. A comparative study of Romance languages*. PhD Dissertation, University of Pennsylvania.
- _____ (1994). "Re-examining negative clauses". In G. Cinque et alii (eds.) *Paths Towards Universal Grammar*. Georgetown: Georgetown University Press, 427-451.
- _____ (1997). "Negation and Verb Movement". In L. Haegeman (ed.) *The new comparative syntax*. London: Longman, 214-245.
- _____ (2001). "Sentential negation". In M. Baltin & C. Collins (eds.) *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell, 511-535.
- Zeijlstra, H. (2004). *Sentential negation and negative concord*. PhD Dissertation, University of Amsterdam.
- Zubizarreta, M. (1998). *Prosody, focus and word order*. Cambridge: The MIT Press.