Tense features and argument structure in Capeverdean predicates

Doctoral dissertation presented to
Faculdade de Ciências Sociais e Humanas
Universidade Nova de Lisboa

Author: Fernanda Pratas
Supervisor: João Costa

Finished: December 2006
Defended: May 2007

Juri: Michel DeGraff (MIT), Inês Duarte (Universidade de Lisboa), Clara Nunes Correia, João Costa, Maria Lobo, Ana Madeira, Isabel Tomás and Maria Francisca Xavier (UNL).
Table of contents

Acknowledgments........................................................................................................ Erro! Marcador não definido.

Chapter one

Introduction – what’s in a verb?............................................................... 1
1.1 Goals .................................................................................................................. 2
  1.1.1 Previous problems and present questions......................................................... 2
  1.1.2 Summary of questions....................................................................................... 7
1.2 Framework............................................................................................................ 9
  1.2.1 Some notes on Government & Binding............................................................. 11
  1.2.2 Some notes on Minimalism............................................................................... 12
  1.2.3 A somewhat intertwined perspective............................................................... 14
1.3 Creole or Capeverdean? “The name of the rose”........................................ 17
1.4 Practical issues on the field............................................................................ 20
1.5 Contents description....................................................................................... 28

Chapter two

Sentence structure ..................................................................................... 30
2.1 Why not an AgrP node................................................................................... 31
2.2 Are adverbs reliable diagnostics? ............................................................... 33
  2.2.1 Capeverdean restrictions................................................................................ 34
  2.2.2 Meaning, prosodic weight, lexical specifications and categorial status............. 36
2.3 TMAs and other functional units................................................................. 39
  2.3.1 The true value of Aspect................................................................................. 45
    2.3.1.1 Temporal perspective(s): facts for a compositional approach....................... 46
    2.3.1.2 Meaning and structure............................................................................ 59
  2.3.2 Capeverdean TMAs: what tense is it? ............................................................ 61
    2.3.2.1 Ta, sata and the post-verbal -ba................................................................. 63
    2.3.2.2 More on sta and sata.............................................................................. 86
    2.3.2.3 On statives versus non-statives............................................................... 97
    2.3.2.4 The puzzling distribution of dja .............................................................. 107
    2.3.2.5 Some structural clues regarding dja......................................................... 121
    2.3.2.6 Concluding remarks on TMAs............................................................... 122
Acknowledgments

Despite the long weeks of seclusion in the final period of writing and editing, this dissertation includes a great deal of extended team work. In fact, if I dared bring all these people together the result would be more like a crowd than a team. I want to thank them all for their help, data, comments, suggestions and, most specially, for their patience. Nevertheless, all errors and misconceptions in the present text are of my own responsibility.

Some names of this huge list of individuals must be specifically mentioned here.

My ever-lasting gratitude, respect and admiration go to João Costa, my supervisor, for being so wise, open-minded, tireless, human and cheerful. Oh, and yes, so understanding with my childish questions (in a rather mysterious way, he has always managed to keep me taking interesting paths departing from them). I have sent him some of these nervous doubts by e-mail, during my stays in Cape Verde and Cambridge, MA. A few hours later, at most, I would get his attentive replies, tempered with good-humored comments that proved to be so supportive in these tough moments away from home. I shall never forget that.

The second person I must mention is Ana Josefa Cardoso, my most precious Capeverdean informant. Not only is she greatly qualified as an informant (extremely competent, careful in her judgments, sensible to my obsession with the details), she is also very available, patient and vivacious. I must thank her, too, for the essential contacts in the inland of Santiago Island, mainly her family and friends in Flamengos. Mantenha txeu, Josefa!

To other professors in the Centro de Linguística da Universidade Nova de Lisboa I must thank the cooperation and help. My special thanks go to: Maria Francisca Xavier, who has an admirable and enduring encouraging power; Clara Nunes Correia, for her competence and energy and in particular for the help with some Aspect subtleties; and finally Maria Lobo, who read and commented the first draft of this text, pointing out, not only theoretical errors and questions on data, but also my various English misconstructions (so many null subjects “borrowed” from my mother language!!).
In Universidade de Lisboa, I thank especially professor Inês Duarte, who has helped in different crossroads along the way (starting in my Masters dissertation): her wise comments and suggestions have always been of great help.

In my colleagues and friends in the department I have met the most congenial environment. I thank in particular João Loureiro, Ana Castro and especially Maria Alexandra Fiéis; with each of them I have shared research topics on Capeverdean, in the most practical and rewarding sense of teamwork.

To my creolist companions in Portugal and abroad I thank the immediate empathy and the sharing of some specific advice regarding creolist studies. I am especially grateful to Tjerk Hagemeijer and Nélia Alexandre, from Universidade de Lisboa, for the generous and insightful discussions, and also for the cheerful moments we have shared in Lisbon, Coimbra, La Coruña and, with Nélia, in Praia, Cape Verde.

To Marlyse Baptista I address a particular acknowledgment: despite not always sharing the same views or analyses of Capeverdean phenomena, I greatly admire you! Both personally and through your work, you have been a huge source of positive inspiration.

I am indebted to many professors and colleagues in different places. My gratitude goes to them all. I must mention some of them who have spent an amount of their precious time listening to my doubts and giving adequate advice: Alec Marantz, Andres Pablo Salanova, Cristina Schmitt, David Pesetsky, Danny Fox, Jonathan David Bobaljik, Justin Fitzpatrick, Klaus Abels, Michel DeGraff, Morris Halle, Noam Chomsky, Norbert Hornstein, Norvin Richards, Pranav Anand, Richard Larson, among others.

I also thank the members of my jury, who have been able to convert a virtually uncomfortable moment into a pleasant, insightful discussion. The deeply productive consequences of their wise remarks will forever illuminate my work.

During my different periods at MIT I have met the most stimulating work atmosphere; it has indeed been a dream come true. The fact of being away from home, however, would have made things emotionally difficult, had I not count on such good friends as Amelia Reis Silva and Andres P. Salanova, Cristina and Alexandre Ximenes. With all their warmth and generosity they have been able to actually make me feel at home.
My very special gratitude goes, obviously, to all my Capeverdean informants. They are so many that there is a great risk of leaving some name aside. I will take this risk, though. The list is big and spreads over different places. Lisbon: Lalaxu, Euson, Dulcelina and Ana Josefa. Cambridge, MA: Carlos and Jaíza. Flamengos (Santiago Island): Kulau and Socorro, Idalina, Lurdes, Arminda, Nha Domingas, Nha Morgada, and the children Mikeas, Rebeca and Patrícia, among many others. Picos de Santa Catarina (Santiago Island): Deolinda and Jeremias, Édina, Cadí, Elvis, Leno, Nha Marina, Nhô António, Nhô Eduíno, and the children Clever, Eder, Jéssica and Leila. Praia (Santiago Island): Wosvaldu, Ronaldu, Jau, and especially Teresa and Rosalino. I want to thank also the teachers Madalena (high school) and Inês (basic school), the writer Tomé Varela da Silva, and the writer, linguist and now Ministro da Cultura Manuel Veiga.

Finally, I am indebted to my family, who has been of great support in most ways. I also thank my parents, who always believed in my competence and intuitions, and can no longer share these moments with me.

Particularly Adriana, my daughter, I owe daily inputs of energy and wit; she has managed to pass me these feelings even when I was away – how great and warming her e-mails were! She is the liveliest proof that life is beautiful, no matter what.

To Luís, my husband, advisor and best friend, I am eternally grateful for all the marvelous things we share, making everything else seem easier to face. In particular, I thank him for being the primary inspiration of my work on Capeverdean, and for all the practical assistance and support.

The four years of research, in Portugal and abroad, that led to the present dissertation have been funded by Fundação para a Ciência e a Tecnologia, under the reference: SFRH / BD / 11064 / 2002
Chapter one

Introduction – what’s in a verb?

This dissertation is about verbs, particularly Capeverdean verbs. The vagueness of this statement is intentional, for such broad a scope comprehends most topics, if not all, that will be addressed, described and discussed over the next three hundred pages. These specific topics are distributed over five main chapters – from Two to Six.

Verbs share the typical properties of the protagonists in a complex novel. They dominate every episode. Even though some of them are weak, and their presence would not bring any relevant contribution, we may miss them (“Me Jane, you Tarzan”). All the details in their environment submit to their internal idiosyncrasies: the number of obligatory arguments in a sentence; the category of these arguments – DPs, PPs or CPs – and the theta-roles assigned to them; which Case features must be checked in the derivation; what type of markers for Tense, Modality and Aspect are needed for a certain reading to obtain; etc. When some of these environmental details show some strong personality (like the language specific syntactic requirement that indirect object and direct object appear in a certain order – or that they display a given hierarchic structure), verbs do not surrender, imposing then some restrictions of their own (I sent Mary a letter / I sent a letter to Mary vs. * I sent America a letter / I sent a letter to America).

These protagonist-like properties of verbs (not to mention right away the huge diversity that these linguistic heroes may show) make this dissertation the “natural” following step to my Master’s Degree, accomplished in May 2002. The work developed then was on the pronominals of Capeverdean (Santiago Island variety). In the present Chapter, section 1.1 describes the passage from pronominals to verbs. Section 1.2 contains a sketch of the theoretical framework that has guided my recent research and the elaboration of the present text. A short but hopefully clear explanation for the reason why the mother language of Cape Verdeans is referred to as Capeverdean, as well as
some remarks on the compartmentalization of the so-called varieties, appears in section 1.3, and section 1.4 presents a synthesis of some relevant observations concerning one of the most cherished parts of linguistic research: fieldwork. Finally, section 1.5 summarizes the contents of the different chapters.

### 1.1 Goals

Capeverdean is a Creole language scarcely studied within a Generative Grammar framework. Even if we consider only the variety spoken in Santiago, the first island in the archipelago to be populated by Portuguese and the slaves brought by these from the African Continent, back in the fifteenth century, the language presents such a huge amount of linguistic material to be gathered, organized and analyzed that, on one side, it offers a stimulating diversity of topics to pursue, and on the other side this diversity can cause confusion and some important time losses.

As a necessary strategy, it seemed reasonable to pick up some of the unsolved problems of my Master’s dissertation in order to pursue some lines of inquiry established then. In the following subsections I will summarize the problems left open in my previous work (1.1.1) and formulate the questions that will guide this dissertation (1.1.2). As we shall see, these questions are oriented to the ultimate goal of this dissertation – a wide, although not exhaustive, description of the Capeverdean verbal system – and they present some clues sustaining the final proposed answers. In the final section of this chapter (1.5), a summary of these tentative answers is available.

The next subsection introduces some of the problems left open in my Master’s dissertation which I have chosen to address with some detail in the present work.

#### 1.1.1 Previous problems and present questions

After having listed the pronominal expressions available and having described their distribution, one of the problems left open concerned argument structure. Some verb entries do not realize any anaphoric expression in reflexive contexts, giving way to puzzling constructions that look like unaccusatives (1):
The analysis presented then made use of the hypothesis, proposed in Reinhart 1996, of a lexical reduction of the internal argument. Following Chierchia 1989, the author contends that some pairs of reflexive/unaccusative entries are derived from a transitive entry. The author specifies a Lexical Uniformity Principle, according to which each verbal concept corresponds to a lexical entry with a specific thematic structure; the different thematic forms of the same concept are derived through lexical operations. Some of these operations have been considered in the literature (for instance Williams 1985 and Grimshaw 1990), but they can be synthesized in two: one that operates on the external argument, resulting in passivization; and one that operates on the internal argument, resulting in reflexivization. This last one – reduction of the internal argument – would be the one at stake in (1).

This approach allows for a closer analysis of Burzio’s Generalization² and of the interaction between derivational, lexical and morphological properties, which the author tests for four different languages: Hebrew, Dutch, English and Italian. One of the flaws of this hypothesis, however, is that it leaves unsolved, for instance, the question of what happens to Accusative Case.

Some of the other previously unsolved problems are about subjects.

Namely:

i) the status of the language with respect to the null-subject parameter – given that it allows no null referential subjects in matrix contexts (2b) and also that, since there is no expletive pronominal available in its lexicon, it does not display any phonological subject in any type of “expletive” constructions(2c);

---

¹ For the moment we can say, rather simply, that the absence of any TMA marker generally means a past perfect reading for Capeverdean non-stative verbs.

² Burzio’s generalization goes as follows: All and only the verbs that can assign a thematic role (also called theta- or θ-role) to the subject can assign accusative Case to an object (Burzio 1986:178).
ii) the nature of the empty category available in certain contexts that look like control (2d); if there is a PRO, what is its distribution, and if there is not, what strategies can be considered for the analysis of these contexts.

(2) a. *N sta duenti.
   ‘I am sick.’
b. *Sta duenti.
c. Txobi txeu.
   rains a lot
   ‘It rains a lot.’
d. *N fla Maria p’-e ben djanta. / N fla Maria pa ben djanta.
   I told Maria PREP-3SG come dinner
   ‘I told Maria to come to dinner.’

At the time, I considered Capeverdean as being a semi-pro-drop language, leaving the problem of control constructions for further investigation. But this further research led to the questioning of the nature of expletive *pro and referential *pro as being parts of the same phenomenon.

Even further, other big questions arose:

i) what exactly is a null expletive / expletive *pro?
ii) why do languages need it, since it is an entity which has neither phonological nor semantic content?
iii) more specifically, does Capeverdean need an expletive *pro?
iv) what if ‘no subject’ means that, at least in some languages, the subject position is not projected?

A different problem, although also possible to relate to subject properties in various ways that I hope will become clear over the next chapters, concerns the functional domain. I have proposed – and will sustain – that Capeverdean subjects move from their internal VP positions (Spec,VP or complement of V, depending on the verb being transitive or unaccusative) to Spec,TP. The proposal that Spec,TP is the functional
landing site for DP subjects follows from the lack of empirical evidence for considering an AgrP node, since both possible lines of argumentation in favor of this category do not hold in the language.

In other words:

i) do we really need to consider an AgrP node if:
   a) there are no positional arguments motivating this projection, and
   b) the language displays no overt verbal agreement morphology?

ii) are there some other morphological / abstract features involved?

At the time, some other questions on the structure of the functional domain have been addressed. Consider the list of sentences in (3):

(3) a. *N ta kume karne.*
   1SG TMA eat meat
   ‘I eat meat.’

b. *N ka kume karne.*
   1SG NEG eat meat
   ‘I didn’t eat meat.’

c. *N ka ta kume karne.*
   1SG NEG TMA eat meat
   ‘I don’t eat meat.’ / ‘I am not going to eat meat.’

d. *N sata kume karne.*
   1SG TMA eat meat
   ‘I am eating meat. (at this very moment / at this period of my life)’

e. *N ka sata kume karne.*
   1SG NEG TMA eat meat
   ‘I am not eating meat. (at this very moment / at this period of my life)’

And also the contrasts in (4):

(4) a. *N ta gostaba di bai / baba Portugal.*
   1SG TMA like TMA of go go TMA Portugal
   ‘I would like to go / have gone to Portugal.’
b. N   gosta   di mora   na Portugal.
   1SG  like   of live   in Portugal
   ‘I like living in Portugal.’

c. N   sata   gostaba   di mora   na Portugal
   1SG  TMA  like.TMA   of live   in Portugal
   ‘I was loving to live in Portugal.’

And in (5):

(5) a. Sata txobe oxi.
   TMA rain today
   ‘It is raining today.’  [also possible: ‘It is (certainly) going to rain today.’]

b. Oxi sata txobe.
   today TMA rain
   ‘It is raining today.’  [??’It is going to rain today.’]

Some of the questions that are raised at first glance were left unanswered:

i) how can we map onto syntax the different morphemes available in the language for marking Tense, Modality and Aspect?

ii) in other words, can we actually consider that each distinct morpheme marks a specific value (even though they may be altered under a compositional perspective), so that we can safely adopt a label for each distinct projection where they surface on?

iii) what are the specific idiosyncrasies of verbs usually classified as statives, such as gosta ‘like’, as opposed to the ones classified as non-statatives, such as kume ‘eat’, regarding the selection of different morphemes to mark each desired temporal value?

iv) what about expletive constructions, such as the ones with weather predicates like txobe ‘rain’? do they have demands of their own?

v) which relation with abstract features do functional projections and/or the various classes of predicates bear? for instance, does the main functional head in a construction with a verb that does not select for
an external argument (subject) still bear a Case feature searching for an agreement relation with some NP?

vi) which positional relation do TMA morphemes bear with other functional materials, such as negation and subject clitics?

At a certain point it became clear that this bundle of unanswered questions could be organized under one label – the verbal system of Capeverdean.

1.1.2 Summary of questions

The order of the questions at stake in this dissertation does not follow the order in which the correspondent open problems have been presented in the previous subsection. The reason is the following: 1.1.1 has been organized in the logical (and chronological) succession in which the problems arose; the present subsection projects in thematic groups (identified by A, B and C) the questions that will guide the present work, although not necessarily in the same order. The intention is that they be addressed, all of them, over the different chapters and sections, whenever it seems appropriate, but with no correspondence to any specific order of their own.

The main topics, and respective questions, are the following.

A. Clause architecture:

i) Which functional categories, among the immense array that has been proposed in the literature in the last decades, seem to fit the language, being able to accommodate (if any) syntactic/abstract Tense features, morphological markers for TMA, as well as negation?

ii) Is there a universal list of functional projections accommodating the necessities of all languages? Or, on the contrary, is there evidence to follow the proposal of a Split, IP parameter (Bobaljik 1995, a.o.), according to which some languages have the IP domain split in different functional nodes (namely an AgrP node) and other languages do not?
iii) How is it possible to make room for the different markers that help build the temporal-aspectual meaning of the sentence, if we make the option, grounded on an empirical basis, of not considering an array of functional categories?

B. Semantic properties of predicates:

iv) What are the properties of the so-called stative (sabe ‘know’) vs. non-stative (kanta ‘sing’) types of predicates in the language? What are the consequences of these properties for the selection of certain distinct markers for Tense, Mood and Aspect, in some given specific contexts?

v) Which specific aspectual properties may be considered inherent to some particular verbs?

vi) What is the relevance of the traditional verb typology transitive / intransitive / unaccusative, supported by observations on argument structure and thematic role assignment? In which way can this help us deal with phenomena such as the apparent unaccusative structures that have a real reflexive reading?

vii) What is the role of the theta-criterion when we face some challenges such as an NP which seems to bare two distinct thematic roles?

C. Purely syntactic requirements:

viii) Is there a reason to sustain the stipulation that all sentences must have subjects, even though some sentences have neither an external nor an internal argument to get a subject theta-role?

ix) What is the nature of the so-called expletive pro, which seemed empirically obligatory in Capeverdean in spite of the prohibition of referential pro in most embedded contexts and in all matrix contexts? Does the language need this empty category?

x) What is the nature of PRO, the empty category which traditionally occupies the subject position of non-finite clauses, and which has been given a very special treatment in the theory, with a theorem and a module of its own? What lies beneath some
apparent control contexts in Capeverdean, given that in some cases the supposed controlled embedded subject of an infinitive is obligatorily overt (N fla Maria p’-e laba. / * N fla Maria pa laba.
– ‘I told Maria for *(she) to wash [herself].’

xi) What is the nature of the syntactic entity that has been used to support this stipulation: the puzzling EPP (Extended Projection Principle), which began as the requirement that all sentences have subjects (Chomsky 1982)? It has been reformulated as a head feature (Chomsky 1995) and, lately, as the property of a feature of a head. What do these new formulations bring to the analysis of Capeverdean phenomena?

Some of these questions are more general and intertwined, and it becomes easily clear why they can be organized under the same topic. Some are more specific but still connected to others, although the way to figure out this connection is a little more indirect. And some are so particular that to include them here may seem a statement of a somewhat tortuous line of reasoning.

The greatest challenge I assume in this task of bringing them together is also the simplest and minimal one: to make any sense.

In section 1.5, the last part of this Introduction, I will outline the distribution of these questions along the different chapters of this dissertation.

1.2 Framework

This section presents a synthesis of the theoretical approach adopted, pointing out some defining aspects of the expressions Generative Grammar, Principles-and-Parameters, Government and Binding, and finally Minimalist Program.

Generative Grammar assumes a simple answer to grammatical theory’s central problem. One possible enunciation of this problem, in some contexts referred to as the Logical Problem of Language Acquisition (LPLA), runs as follows: how is it possible that children are able to acquire, so early and so fast, such a specific and complex grammatical competence given the impoverished nature of the available data? The
simple answer from a Generative Grammar perspective is this: all healthy children are endowed with an innate biological language device, consisting of a set of principles of grammar. This set of principles, the Universal Grammar (UG), comes along with some open parameters, for which the child will have to set a value according to the primary linguistic data, i.e., the input provided by the specific language(s) spoken around her. We must note here that this input:

i) occurs early (from the first days of the child’s life, or even during conception);
ii) is consistent and redundant (speaking adults use parameter values already set for their mother language);
iii) is very limited in quantity, especially if we compare it with the richness and complexity of the child’s linguistic knowledge at very early stages of acquisition.

Once all parameter values are set, the child has acquired the specific grammar of her native language. This is an elegant approach to the apparent contradiction between an idea of Universal Grammar and the grammatical specificity of particular languages.

Government and Binding (GB), launched by Chomsky in the 1980’s (Chomsky 1981, 1982, 1986 and 1986b), is the best-known theory trying to demonstrate the viability of such an approach. I will describe its relevant features in subsection 1.2.1.

In the following decade some concerns were added to the search for a scientific solution to grammatical theory’s central problem. These concerns could be synthesized as a sort of a lemma: less can be more. From Chomsky 1993, the approach known as Minimalist Program (MP) has developed and reinforced the notion of “economy”, applying it to two distinct plans: the methodological plan (speaking of theories, the least the best: where we can account for some phenomenon with the description of one relation or one level, for instance, we should refrain from doing it using two or more of these descriptive tools) and the linguistic plan (speaking of languages, grammars are designed / conceived in a way that maximizes resources, obtaining the best results with the least effort). I will describe the relevant features of MP in subsection 1.2.2.

This research program – which is not a theory, but rather a bundle of theoretical concerns (and, as can be seen in the current theoretical debates, the translation of these concerns into practical methods is not always easy) – has discarded some of the
important devices achieved by GB. Since then, however, many linguists have found it appropriate not to follow these rejections blindfolded. The alternative is thus to assume the general idea of minimalism – for also in other scientific fields, like Mathematics and Physics, there are strong concerns of elegance and simplicity –, but without abandoning the explanatory adequacy of some GB formulations. This dissertation will focus on this mixed approach, and I outline the main reasons for this option in subsection 1.2.3.

1.2.1 Some notes on Government & Binding

The relevant features of GB, as a theory developed from the principles and parameters approach, are the following (mostly adapted from Epstein & Hornstein 1999 and from Costa & Figueiredo Silva 2005):

- the grammar is divided into different subcomponents (modules);
- there are four different levels: D(eep)-Structure (DS); S(urface)-Structure (SS); Logical Form (LF) and Phonological Form (PF);
- despite Chomsky’s (1986b:93) characterization of GB as a “virtually rule-free system”, GB has in fact some rules: e.g. Case assignment rules, which have the property of applying only in government configurations;
- we must note here that Case assignment may occur under a specifier-head agreement relation as well; in some languages, both options for Case assignment are available (this is one of the cases of redundancy to which MP will be so sensitive);
- Move-α is also a GB rule – it is the basic movement rule, and it allows any category to move anywhere at any time (movement is not driven by any special motivation); this unconstrained transformational component gives way to an overgeneration of products; among these, the ones that do not satisfy various conditions of well-formedness are then filtered out at different levels (the level where some structure is rejected depends on the nature of the violated requirement);
- government is the grammatical relation that guarantees formal unity to this complex scenario.

Some more crucial concepts in GB will be mentioned in the next subsection, along with the new, minimalist, terms that have replaced them.
1.2.2 Some notes on Minimalism

The breakthrough mark of MP, a theoretical proposal starting in Chomsky 1993, has been to frontally assume the biggest of all questions: why?

Why do these (GB’s or any other) stipulations hold (if they hold at all)? In order to pursue a possible answer we must follow a basic scientific effort: any stipulation must be logically deduced (and this deduction must be formally expressed) from a minimal number of theoretical premises.

Having this main purpose in mind, MP developed some theoretical devices: some of them are more generally satisfying, some others are more problematic. Among these problematic ones, some have been successively replaced and/or modified by MP theoreticians, and are still under discussion. The most relevant of MP features can be listed as follows (adapted from Epstein & Hornstein 1999, from Epstein & Selly 2002 and from Costa & Figueiredo Silva 2005):

- principles and filters are no longer determinant; what is in charge here is some kind of balance between rules on one side and constraints on the application of these rules on the other (while GB intended to be a “virtually rule-free system”);
- there are only two levels of grammar, PF and LF; in other words, language is sound and meaning; the reduction in the number of levels is an evident choice obeying a desired principle of parsimony (two are better than four);
- Lexical Items (LIs) have no features other than those interpreted at the interface (properties of sound and meaning); lexical features are the atomic elements upon which operations compute;
- the grammatical module interacts with components that interpret sentences at the phonetic and the conceptual interfaces; these interfaces that meet PF and LF are two systems not entirely linguistic: respectively the articulatory-phonetic and the conceptual-intentional; grammatical objects not fully interpretable at these interfaces cause the derivation to crash;
- recursion, a property of human language that allows sentences to have no length limit, obtains by transformations (it is no longer a property of DS; DS no longer even exists);
- Merge is the basic grammatical operation: external Merge being the one that builds sentences from words and internal Merge (or Move) being the one that accommodates phenomena of displacement;
- Move does not apply to lexical items but to features, F – that is, the operation Move-$\alpha$ is replaced by Move-$F$; the operation Move is driven by the morphological requirement of a feature to be checked;

- Move is a last resort operation (F can raise only if a checking relation, between F and a sublabel of K, or between F and K itself) is established; features that are uninterpretable at one of the interfaces (like Case, for instance) must be checked (this is the way of eliminating them) in order to prevent the violation of full interpretation; note that Case is no longer “assigned” but “checked”\(^3\); morphological features like Case and agreement are checked in functional domains; under economy requirements, movement must be driven, and the need for an uninterpretable feature to be checked is the only licit reason for Move to occur;

- a checked feature is deleted – rendered invisible at LF but still accessible to computation – when possible; a deleted feature is erased – it becomes inaccessible to computation – when possible;

- operations like Move and Attract $F$ apply depending on the features that are detected when lexical items are scanned. There are semantic, phonological and formal features; they are also intrinsic or optional, weak or strong, and plus or minus interpretable.

- that is, from the uninterpretable and interpretable features, some are weak and some are strong; strong features must be checked in overt syntax, while weak features must be checked in covert syntax; this distinction accommodates the difference between, for instance, English and Chinese regarding wh expressions (these expressions occur in situ in Chinese, where the relevant features are checked in the covert component – after Spell-Out –, and move in English, where the relevant features are checked in the overt component – before Spell-Out); this relation between the strength of features and the need for them to be checked in different components reflects a minimalist approach to control, to the EPP and to pro-drop properties, among others;

- cyclicity is the basic notion for the iterative recursive application of a rule;

- the length of displacements is a very important matter: Move must be short; so, when the relevant features involved in the operation – movement is the result of one feature attracting another – are substantially apart from each other, some cyclicity

\(^3\) Epstein & Selly 2006:34 suggest that Case may be a PF uninterpretable feature, and not an LF uninterpretable feature.
notions are introduced to avoid a movement excessively long (Relativized Minimality is a similar notion already introduced by Rizzi 1990);

- features are checked locally, i.e., at the level of a phrase structure (Move occurs so that these locality conditions are satisfied): a strong uninterpretable feature targets the specifier of the head containing the interpretable feature that attracts it, and in this case the whole category must overtly move along, resulting in a specifier-head configuration; a weak uninterpretable feature is attracted by its interpretable counterpart but it is subject to covert movement (there is no requirement for the whole category to move along); this weak uninterpretable feature is thus checked in a head position adjoined to the target head;

- no new features are introduced by the computational system of human language (ChL) – this is the Inclusiveness Condition, which means the prohibition of all nonexplanatory representational encodings: subscripts, superscripts, bar-levels, and also chains (these used to be defined on S-structure representations) and star-marking (this star-feature has been used in Chomsky & Lasnik 1993 analysis);

- thematic roles (theta- or $\theta$-roles) are not morphological features, like Case; $\theta$-role assignment takes place in the lexical domain and it cannot license movement.

### 1.2.3 A somewhat intertwined perspective

The intuition that the general MP conceptual proposals are on the right track does not preclude some particular questions. The most primary of them is:

i) should we throw away all those GB achievements that seem undesirable under MP assumptions?

And some more precise ones:

ii) what exactly do we need for a stipulation to be logically deduced (and to formally express this deduction) from a minimal number of theoretical premises?

iii) what does “minimal” mean?

iv) what does “economy” mean?
In their “Introduction”, Epstein & Selly 2002 quote American physicist Richard Feynman: “[…] the real glory of science is that we can find a way of thinking such that the law is evident.” But when can we be sure that some law is evident? They also quote Albert Einstein: “Can we hope to be guided by experience at all when there exist theories (such as classical mechanics) which to a large extent do justice to experience, without getting to the root of the matter? I answer without hesitation that there is, in my opinion, a right way, and that we are capable of finding it. Our experience hitherto justifies us in believing that nature is the realization of the simplest conceivable mathematical ideas.” Even though we are not sure about the meaning of words like “simplest”, we can assume a trustful attitude and follow the genius belief.

This is now the moment to build a path from these beautiful conceptual observations to the practice of linguistic research. Given the above descriptions of GB, firstly we must decide whether we really find all MP concepts as actually new or rather similar to some of these old concepts, only formulated in a different way. Supposing they are truly different tools, we also need to figure out whether we prefer the old stipulations and their descriptive power, or choose to simply discard them because they are not:

i) evident laws;
ii) answers to the question “why”, or
iii) obedient to properties such as simplicity and minimality.

Let us assume for a moment that this last one – the rejection of GB – would be our choice, for we think economy concerns are the only healthy way to guide our research steps. What would then replace the discarded GB stipulations? Would it be MP concepts and stipulations? First of all, the aspect to be assured must be that these new concepts offer new achievements – in the sense of more clear explanations and more adequate generalizations – than the old ones.

According to Chomsky, “Minimalist demands at least have the merit of sharpening the question of whether we have a genuine explanation or a restatement of a problem in other terms.”(Chomsky 1995:233-4) Ironically this seems to be the problem with some of the MP ideas: they aim to give a genuine explanation, for this is the genuine conceptual compromise of Minimalism, but they are accused of being nothing or little
more than a “restatement of a problem in other terms.” Some novel notions have been introduced, namely Agree (a mechanism under which features can be checked without movement) and Phase (to help define the domains in which Agree can occur and to establish the cyclicity of the derivation), about which it is legitimate to ask: are they that much different from the old notions of, respectively, Government and Barrier?

On the other side, there are other new notions that really bring some considerable advances.

One example of these new notions is the one assuming that some properties at the level of the interfaces must be addressed at these same levels, a wise distribution that allows sparing much machinery at the syntax/computation level. This is the position assumed, for instance, in Costa 2004, supported by the author’s analyses of European Portuguese word order phenomena: “the computational system generates multiple convergent outputs; interface constraints may filter or select them.” This means that “syntax proper may be reduced to its own tools, not having to manipulate semantic, discourse or prosodic variables.” (Costa 2004:160)

Another example of these welcome new notions is the copy theory of movement: defining movement as the composite of two more primitive relations, Copy and Merge, gives the operation Merge the ability of covering both local relations (like the coupling of heads and phrases) and long distance ones (like the coupling of antecedents and anaphors). This brings new possible analysis of reconstruction phenomena, among others. As proposed in Boeckx, Hornstein & Nunes 2005, this notion of movement will even mediate longer distance relations (operator-variable binding, control and reflexivization) by making copies of the relevant expressions and merging them in the required phrasal positions, deleting, or not, the original expression.

As for the notion of parameter, in GB we had the typical Null Subject Parameter (NSP), also named pro-drop parameter, submitted to different definitions and revisions and still rather confusing after all these years. In the view of MP, possible parameters are features (and their values, such as weak/strong) of functional categories, which allows for a relation between various phenomena. In other words, for instance a strong feature on T in a given language (while it is weak in others) can account for facts regarding null subjects and verb movement. The next big problem, however, comes with the need to define a whole array of relevant features and, moreover, to decide which of them are strong (a trigger for visible movement, since they must be checked
before Spell-Out) and which of them are weak (a possible trigger for non-visible movement, since they may be checked after Spell-Out).

The escape from this trap is contained in the MP itself: this is a research program, not an ultimate theory.

Given all this, I can say by now that the research results and tentative analyses included in this dissertation have been guided by minimalist concerns of economy and simplicity. Nevertheless, in order to deal with empirical data and the multiple tasks involved in descriptive work, and also in order to present some tentative explanation of some empirical problems, I have used theoretical instruments both from GB and MP eras. The criterion to choose each of them has been, I hope, a legitimate one: the intuition that this chosen instrument would lead (presumably) to a simpler and more evident law.

1.3 Creole or Capeverdean? “The name of the rose”

Since I am not a native Capeverdean Creole speaker, the data under analysis in this dissertation has been gathered in various fieldwork sessions with Cape Verdean informants: these sessions were more intense and systematic during three different periods in Cape Verde (Santiago Island), in September 2001, March 2004 and February 2006, although completed by many other interspersed interviews with Cape Verdean informants mainly in Lisbon but also in Cambridge, MA. Some other pieces of data have been borrowed from other authors’ works, mainly with comparative purposes, and in these cases the sources are locally identified.

At this point, however, there is a specific topic I want to emphasize. One of the most emblematic individuals with respect to linguistic fieldwork has been Ken Hale, from MIT. Writing about his experience, almost twenty years ago, while establishing a research program on Ulwa (Southern Sumu), an indigenous language of the Nicaraguan Atlantic Coast, he gives no way for doubts: “Do whatever you need to do in order to learn the language.” (Hale 2001:81) In spite of the fact that I have tried to follow this advice, and although at some point in my investigation I thought I had some intuitions on the grammaticality of some sentences or expressions, these “feelings” only served as some kind of clue, in order to make decisions on the topics for my fieldwork. Never, ever, have I trusted blindfolded any of these intuitions, and so even those sentences
have been carefully tested with some Capeverdean native informant, either in Cape Verde, in Lisbon or in Cambridge, MA.

For some analysts of all kinds – sociologists, anthropologists, sociolinguists and even politicians – the fact of not being a speaker of a language under analysis is a good reason for not taking any stance on some hot topics under discussion, concerning this same language. In other words, this must be left to the intellectual elites who live in the language native country. At least one of those topics (not the most controversial, though) has to be addressed here: should this language be named Creole, Capeverdean Creole or Capeverdean? The practical reason of having to refer to the language in the present dissertation (and in previous conferences and papers as well), using this or that proper name, made me take a position anyway. I have chosen to name it Capeverdean.

This decision is based on a very clear and certainly non-problematic fact: Capeverdean is shorter than Capeverdean Creole. Just Creole is even shorter, but although in Cape Verde every single person knows what we are talking about whenever we use the word Creole – locally, this is the proper name of their native language –, in some Portuguese contexts people could understand it as Guinea-Bissau Creole, for instance, since this is the native language of people born in another former Portuguese colony. Not unexpectedly, the native population of Guinea-Bissau also names their respective mother language simply as Creole. This would be even more confusing in a broader context, like the United States. Which Creole would we be referring to? It could be Capeverdean (a Portuguese based Creole) or Haitian (a French based Creole), just to name two evident possibilities.

Moreover, Dejean 2001, in a discussion around the question of (not) changing the “name of the rose” in the case of Haitian Creole – which, not surprisingly, is also known by its native speakers simply as Creole –, makes some kind of joke on previous arguments in favor of naming it Haitian, like the ones in Hyppolite 1978 or in Freeman & Laguerre 1996. Dejean 2001 considers the position of Freeman, in Freeman and Laguerre 1996, as “individual, arbitrary and capricious”, and observes that the argument of following a logical relation between the names of the countries and the names of languages spoken by their populations – Italy: Italian; France: French, etc. – is not convincing, since there is no relation between the names of, for instance, Australia or United States and the language spoken by their populations: English. The same non-relation is visible between Morocco or Algeria and Arabic; and the same between Peru or Argentina and Spanish.
In my opinion, Dejean 2001 is comparing totally different things: the English spoken in the USA, the Arabic spoken in Morocco or Algeria, or the Spanish spoken in Peru or Argentina, are varieties of the colonial languages introduced by the people who once imposed their rule leading to all sorts of arbitrary decisions, including what the official language would be. It seems important to underline here that the question is not to suggest the variety of Portuguese spoken in Portugal former colonies – e.g. Cape Verde, Angola, Mozambique – should be named Capeverdean, Angolan or Mozambican. The question here is on the native language for every single person born in Cape Verde (the country has around 400,000 inhabitants) and also for most of the estimated one million of Cape Verdeans in the diaspora. This is not any variety of Portuguese. It is a distinct language, with its own internal variation phenomena, distributed by the different islands and migrant communities. Its name is Capeverdean Creole. Or, for a shorter version which can not be mistaken for any other language: Capeverdean.

At this point it is also crucial to clarify some other terminological details that can be implied on the theoretical approach assumed here, and, as a consequence, also on the ones not assumed. The core of the present work is not an investigation on the emergence of Creoles, in particular Capeverdean, under a diachronic perspective, nor is it particularly focused on the understanding of how some ancient linguistic facts (from substratum languages in West Africa or the superstratum language, European Portuguese) determined such word or such phenomenon. Nevertheless, whenever necessary I will appeal to some interesting achievements that have been presented under the label of creolist studies.

The main purpose in my previous works has been to describe and analyze Capeverdean phenomena from a synchronic perspective, such as many linguists do with, say, English or Portuguese phenomena without necessarily mention ancient forms of Old English or Latin. Given that Capeverdean data have been so scarcely studied under a generative approach, taking some steps forward in this domain seemed to me sufficient as a starting point. Moreover, some creolist studies I had contact with seemed to me more descriptive, with no tentative analyses of the underlying phenomena of grammatical rules and restrictions. Sometimes, some of them seemed also contaminated by the kind of social perspective that I strongly repudiate: the traditional vision of Creoles as
languages spoken by people living in exotic places and (similarly to their colorful clothes, their different complexion and their cultural rituals), so amazing and curious.

Over the years, though, a different angle on the subject has been revealed. Above everything, I have been led to admit that some diachronic perspective can positively illuminate the description and analysis of contemporary data, mainly if we are working with a Creole language. I hope that this late learning has been on time to help me handle with some of the phenomena I intend to bring to discussion in this dissertation. As for the traditional, neo-colonialist, attitude towards Creoles, happily it does not actually hold for most creolists. Moreover, the more creolist studies come to see the light of the day, the more intelligent and scientific grounded proposals and perspectives are likely to appear. The rest will emerge on its own, with time and the necessary persistence.

By “the rest” I mean the results that this will produce in the lives of the people who are native speakers of these languages, and that have hitherto been submitted to a school education in a foreign language – that is, to be taught such early and basic skills as how to read and write, or even the basic mathematical operations, in a language that was completely unknown to them until the day they put step on school. For we must never forget: “In carrying out field research, linguists are inevitably responsible to the larger human community which its results could affect.” (Hale 2001: 76)

1.4 Practical issues on the field

My fieldtrips to Cape Verde have mainly concentrated on the variety traditionally called “deep Creole” – or, in their own language, kriolu txambu –, spoken in the inland of Santiago Island. Namely in the municipality of Santa Catarina, which includes the town of Assomada and the village of Picos de Santa Catarina, where I have been doing fieldwork for my Mestrado dissertation, in September 2001, and also in the municipality of Calheta de São Miguel, namely the village of Flamengos, where I have been in March 2004. In the fieldwork period accomplished in February 2006, with the main goal of confirming some data and also gather new occurrences of some mysterious entities, like the word dja, complex predicates and interactions between different verbs and TMA morphemes, I have not been so specific about the geographical choices. This
time I have interviewed mainly informants in the city of Praia, ensuring that they have had regular contact with the inland more “deep creole”.

A crucial note on the so-called varieties is necessary at this point. Having consistently checked with my informants in Praia some of the previous data obtained in work sessions with inland informants, I have verified that the supposed dialectal differences (between the inland and the capital of Santiago Island) do not interfere with the issues I am dealing with. In some cases, this comparison only gives more strength to the hypothesis that seemed to apply to the inland data. One crucial example, in my opinion, is the puzzling case of embedded subjects in certain clauses which seem non-finite environments; these subjects can be “silent” (to call them “null” already implies a classification, and that is not the point here) with certain transitive entries and must be “heard” with certain reflexive entries where there is no overt reflexive.

(6) a. N fla Djon p’-e / pa laba losa.  
1SG tell Djon PREP-3SG / PREP wash dishes  
‘I told Djon to wash the dishes.’

b. N fla Djon p’-e / *pa laba.  
1SG tell Djon PREP-3SG / PREP wash  
‘I told Djon to wash (himself).’

My city of Praia informants also consistently reject the form in (6b), with a “silent” subject in the embedded reflexive construction. Given that, for the obvious reasons, these informants are more susceptible to interferences from Portuguese, and given that Portuguese allows a “silent” embedded subject in these contexts (although, crucially, these reflexive contexts in Portuguese display the reflexive morpheme se), we could perhaps expect here a case of interference as well; that is, we could expect a possible “silent” embedded subject in (6b). This does not happen, though. Capeverdean data are very consistent with respect to these embedded contexts, which I will address in detail in Chapters Four and Six.

Besides these periods in the island, I have also had numerous work sessions with Capeverdean informants in Lisbon and also in Cambridge, MA. All the informants that have collaborated with me out of Cape Verde were born in one of these inland Santiago
locals that I have just mentioned. Whenever it seems relevant, I will include – and properly identify – data provided by speakers of different places.

The nature of the differences between these proclaimed varieties, however, has not yet been extensively described, although some lexical and phonological differences have been identified (the variety spoken in São Vicente, for instance, has many clear interferences from English – a great number of English families migrated to the island in the XIX century). As for the syntactic differences, if any, they have not attracted much attention in the past.

Pereira 2006 gathers a substantial range of evidence – historical, geographical and linguistic – against the idea that there are two (or even more, according to some views) different Creole languages in the archipelago. She contends that, as any other language, Capeverdean has a sub-set of possible grammars, connected to each other by a common history, by a bundle of invariants and by a set of variants heterogeneously distributed over the different islands. She quotes other authors that have met some difficulty when trying to organize the empirical data into two distinct geographic areas: the Sotavento and the Barlavento groups of islands (there are nine inhabited islands4), which have traditionally been considered, respectively, the more conservative, more marked by African influence, and the less conservative, more akin to Portuguese.

The work in Pereira 2006 is, thus, oriented to the deconstruction of such a myth – of two distinct and homogeneous language varieties –, which has actually been supported by a strong social and political tradition. The fact of there being different varieties has been functioning as one of the peripheral “excuses” for the postponement of the decision of making Capeverdean an official language in the country, along with Portuguese (which has been the only official language up to these days, in a postcolonial situation5). There have been some rivalries mainly between Santiago and São Vicente, which represent respectively the Sotavento and the Barlavento groups of islands. Although, for the sake of clarity, it is fundamental to identify specifically the

---

4 For an extensive description of the geography and history of Cape Verde, see Batalha 2004 and references therein.

5 The other “excuse” has been the more or less open consideration that the “creole” nature of the language implies it is not endowed with the complexities and formal properties needed for a language to assure all the tasks involved in an official status (writing of laws and official documents, official school manuals, etc.). Both these “excuses” might be interpreted as clearly functioning in favor of certain members of an intellectual elite. Having a privileged access to instruction abroad, one might easily feel free to consider that the only true mother tongue in the country is ideal for traditional storytelling, but not for studying sciences or mathematics.
geographic origin of my informants, I will keep away from this debate, leaving it for sociolinguists.

In order to plan the fieldwork sessions which have fed the present work, linguistic facts registered by previous authors have been fundamental. From different perspectives, they all described the grammar of the language: Lopes da Silva 1957, Almada 1961, Silva 1985, 1990, Veiga 1995, Quint 1998, Lang 2002. The most important of these sources, though, has been Baptista 2002, within a Generative Grammar framework.

There has been a period in which the theoretical studies were privileged by a large number of generativists, who relied mainly on English and some other European languages to pursue their in-depth investigations. But lately, thanks to some charismatic hard-workers with a sort of anthropologist soul, many others have embraced a systematic and theoretically guided gathering of new data from scarcely studied or even from endangered languages. In the latter case, they do it with an exceptional motivation, since they are working under the extreme urgency brought by the feeling that some human patrimony and some possibly determinant clues are at risk of being lost forever.

As said earlier, Capeverdean is the native language of the 400,000 inhabitants of the country, and of the estimated one million living in other countries as well (the so-called Capeverdean diaspora). Hence, we may believe that the language is not at risk of disappearing within the next decades. When spoken abroad it is, as expected, subject to many of the language contact phenomena that have been leading to some well-documented linguistic changes. Capeverdean is, thus, like other Creole languages, a very specific case of contact result. Besides its historic, social and economic origins, in some territories it is continuously in contact with, for instance, Portuguese in Portugal and Cape Verde, English in England and USA, French in France or Dutch in the Netherlands, to name just a few.

Even though Cape Verde is an undemanding place to be at work, there are some difficulties inherent to field research that can be strategically circumvented. For an extensive discussion on various practical problems and pieces of advice regarding fieldwork, there are two manuals which have been precious to me: Vaux & Cooper

---

6 On the genesis of Creoles see mainly DeGraff 1999 and references therein.
1999 and Newman & Ratliff 2001. I am certain that both of them, in different ways, have helped me avoid many distressing difficulties.

Given all the particularities of the relation between this language and its lexifier – which is reflected in the relations between Capeverdean native speakers and myself, as a speaker of the European lexifier –, and also given all the careful steps that must be involved in the gathering of the relevant data from a language that I do not speak fluently, there have been, of course, some crucial episodes in my fieldwork experiences.

The description of the intense and memorable human contact involved in these trips to the field is beyond the goals of the present dissertation (they shall be registered elsewhere). There are, nevertheless, a few practical notes that seem opportune here.

Some of them follow from the recommendations in Vaux & Cooper 1999. In their chapter on the necessary cares regarding acoustic phonetics, the authors tell the story about a bridge in Madras, India, which was initially named after the engineer who built it, a man named Hamilton. “This name contained several sounds and sequences not present in Tamil, the local language, and consequently was interpreted by the local residents as a more familiar sequence [ambuton].” This pronunciation has later been identified with the Tamil word ambattan ‘barber’, and the bridge came to be known as Barber’s Bridge. The story intends to show that people may mishear sounds and words in another language and in the need of interpreting them some wrong conclusions may be accepted.

When some phonological/prosodic phenomena are involved, things get more complicated. This is the case with truncation, where people chop parts of words. And it is also the case with, for instance, the reinterpretation of a diphthong as a vowel. In my first contact with some spontaneous Capeverdean sentences I really felt the danger of misinterpreting some of the sounds and words included in them. The case with the sequence [pobai] has been the most striking, though. It corresponds to E pa bu bai… ‘Is for you go…’ (‘It’s for you to go’) – in Portuguese É para ires…. The tricky part is not on the verb, which can be any (E pa bu faze… ‘It’s for you to do…’, E pa bu ben… ‘It’s for you to come…’), but rather on the various phonological phenomena involved in [po]. We could align a sequence like: pa bu → pa-u → po. Notice that, in the previous written studies, the clitic form -u is the typical object form for 2sg, whereas the subject
clitic is $bu$. Moreover, the sequence [po] is also available for the verb ‘put’, like in $N po$ $livru$ $riba$ $mesa$ ‘I put the book on the table’.

The facts involved in cases like these, which are natural and spontaneous for a native speaker of a given language (to whom nobody has taught the rules and constraints involved and nevertheless is a master of their use), may be puzzling to a foreigner, even when he believes himself to know most of the phenomena at stake.

A different note also taken from the same manual (Vaux & Cooper 1999) which has been crucial as a guiding line in my field experiences is the following: accuracy should take priority over efficiency. In other words, quality is more important than quantity. I was dealing with such delicate matters as verbal paradigms, the building of temporal reference and the various subtleties involved in the possible relations between different verbs in the same sentence. This means a detailed description of the various morphemes and their interactions with different verbs and expressions was needed. As Vaux & Cooper 1999 point out, although the pair of sentences ‘he $is$ crying’ / ‘the man from the village $was$ crying’ might seem a good pair for testing the different verb tenses, other consequences may hide in the change of the subject from a pronominal to a full DP.

The strategy of detailed testing has some risks. One of them is that informants may become really bored in eliciting sessions, with repetitions of sentences with one slight change in each time. Whenever possible, I have followed the advice of taking some further strategic steps in order to circumvent this boredom (by way of avoiding marathon sessions, for instance), since the work can result in bad judgements when the mind of the informant is no longer there. The reader might also feel somewhat bored when confronting the extensive second chapter in the present work, for the very same reasons. This extensive description has been, however, and given the purposes of the present dissertation, difficult to avoid.

In the semantics field, there are some critical topics. One of them, which I have confronted myself, is Aspect marking. But topics such as quantifier scope and the value of some prepositions are also slippery. It takes a lot of careful conversation to be sure, for instance, about the possible ambiguity of sentences like ‘Every book has been read by some student’.

One of the most relevant ingredients of a linguist’s intuition is the precise feeling about some lapse in the communication with the informant. This may happen not only
because one of them has not understood the words (a matter of acoustic phonetics), but also because someone did not get the accurate message.

Doing fieldwork in syntax has different requirements and difficulties. We are willing to know about the informant’s grammaticality judgments instead of the exact meanings of words or morphemes. This might seem easier at first sight, since we can have a long list of sentences and only ask them if they are good or bad sentences in their language. Of course we have to be careful with such sentences as ‘The man bit the dog’; this sentence might cause a rejection on the part of the informant in spite of being perfectly well-formed.

A different danger depends on the possible lack of the relevant context for a sentence to be considered well-formed. In Capeverdian, for instance, if one asks some informants about the simple sentence N ta sabeba! (where we have the stative verb sabe ‘know’, marked by morphology typical of non-statives Past Imperfective), the possible immediate answer can be of nonacceptance (the Past (Imperfective) form for this stative verb is N sabeba ‘I knew’ / ‘I used to know’). When this construction is the consequent in a past conditional environment, however, it is perfectly grammatical, and even the informant may be amazed, since he possibly never noticed that before (Si es kebraba nha kopu N ta sabeba ‘If they broke my glass I would know’).

Things may get truly difficult when the grammatical judgements needed are on such complex topics as binding relations or extraction, for instance. And a pretty serious trouble may arise when searching in some languages for the well-formedness of potentially bizarre questions like ‘Which man did you say your cousin believed had killed what?’ Moreover, sometimes we know that there are two forms for the same sentence and we ask for a judgment on one of them. The informant says it is okay. Then we ask about the other form. The informant may say it is okay, too. When we ask several times for the informant to say the sentence, however, he systematically pronounces one of the forms and never the other.

“When informants invent or embellish data, they are generally aiming to help”, say Vaux & Cooper 1999. I believe that the following story, which happened to me during my last trip to Santiago Island, in February 2006, helps illustrate some of the technical risks on the field. Almost in the end of a session, one apparently good informant provided an unexpected set of data. What made these data seem puzzling was an object
clitic inserted between the verb stem and the postverbal TMA marker -ba: E ta odjamba (‘He used to see me’), E ta oadjamba (‘He used to see you’), N ta odjalba (‘I used to see him’), N ta odfasba (‘I used to see them’). In six years of field investigation on the language I had never heard about such a phenomenon; the only grammatical versions for these sentences have the following shape: ta + V + -ba + free form of the pronominal; like E ta oadjaba mi (see 4.1 for the discussion and one proposal on why the object enclitic is barred in this context).

Seeing my astonishment, this informant felt he could warrant that, although not very common, this exquisite order was possible in varieties spoken in the most remote areas of Santiago Island. When I tried to confirm this with other native people, most of them also reacted as if they had never heard about such a thing, but one told me that this could be from a secluded fishing village on the northwest coast. The problem was that, at the time, I could not go there to pursue this clue: I would be coming back to Portugal the next day. I had no other way out, besides asking some more people, mainly those to whom I always go in case of doubt. Finally one of them came up with a hypothesis: perhaps that informant had been living in Guinea Bissau for some time, which could have brought some influence. He had indeed lived in Guinea Bissau, and in the native language of this other former Portuguese colony the postverbal ba is not a bound morpheme (a suffix), as it is in Capeverdean, but rather a free marker for anteriority. Hence, this could explain why the object clitic would not sound badly between the verb stem and the temporal morpheme – in Guinea Bissau Creole even other categories, such as a DP, are allowed before ba7.

In this section I have presented some observations on the necessary cautions whenever fieldwork is the most important source of linguistic data. For some precious advice and contact with previous experiences, one should consult both Vaux & Cooper 1999 and Newman & Ratliff 2001. I have reported what has been to me an emblematic episode regarding the accuracy of the data provided by otherwise good informants. I will leave a more detailed field report for a future work, concentrating the present dissertation on the description of the Capeverdean verbal system.

The next section describes what will be at issue in each chapter.

7 In Kihm 1994 there are some contexts illustrating this, like the one in (i):

(i) "I kunpra pon ba.
3SG buy bread PAST
‘She had bought bread.’
1.5 Contents description

This last section of the Introduction outlines the contents of this dissertation.

Chapter Two, the longest of them all, will be centered on the Capeverdean inflectional domain, and the proposals assumed in my previous works will be developed. Grounded on some distinct theories on temporal reference and supported by extensively described Capeverdean empirical facts, the main proposal will be that the pieces of inflection surface on multiple heads adjoined to T. Besides accommodating all the TMA morphemes, Capeverdean TP is sufficient to account for the negation head *ka*, the copulas *e* (Present) and *era* (Past) and, finally, subject clitics.

Chapter Three will introduce some issues inside the Verb domain, including the parameter values traditionally connected to verbal features. In the first section the nature of Case and agreement will be discussed, as well as the role of these in the licensing of DPs. The main claim will be that morphological, rather than abstract features are at stake in Capeverdean. The second section will address the verb movement parameter. Taking the crucial argumentation surrounding agreement morphology, which does not exist overtly in the language, the main claim in this section will be that there is no empirical reason to posit V-to-I (in the present case, V-to-T) movement in Capeverdean.

Chapter Four extends the description of phenomena inside the Verb domain, which will be accounted for on the grounds of the proposals assumed so far, namely that the verb stays in its base position. The approach to these topics will require the summary of different joint works developed over the last few years. Besides various papers with João Costa, whose main results are mentioned when the respective topics are under discussion, one with João Loureiro and one with Ana Castro, described in Chapter Three, I have shared one topic with Andres Salanova and two others with Alexandra Fiéis, on distinct puzzles concerning the complement position, that is, internal arguments. The specific puzzles are, respectively: the selection of different object pronominal forms, depending on the presence of postverbal *-ba*; the lack of any anaphoric expression in some reflexive contexts; the structure of ditransitives. Each separate section adds to the summary of these previous works some recent developments of my own that justify their inclusion in the present dissertation, for a
view of the whole puzzle of Capeverdean verbs and their idiosyncrasies has necessarily brought up some adjustments and also new clues to be pursued in the future.

Chapter Five is focused on the null subject parameter. After the analysis of relevant data, it will be argued that there is no empirical reason to state that Capeverdean is a pro-drop language: referential pro is prohibited in all matrix clauses. In expletive constructions no expletive subject is overtly realized, which, together with the proposals in the previous chapters, will lead to the conclusion that, when not needed, Spec,TP is not projected. In order to account for this, the Extended Projection Principle (EPP) will be discussed, ending up in the proposal that in Capeverdean the EPP does not hold. In other words, T does not have a second edge feature in the language: its specifier is only projected whenever some element must move to this position (internal merge).

Chapter Six will deal with embedded non-finite clauses. This will call into discussion a good part of what has been at issue in the previous chapters, namely argument structure and Tense features. First of all, given that there is no specific morphology for infinitives in Capeverdean, some possible diagnostics will be essayed. Once infinitives are identified, some hypothesis on its internal structure will be proposed, also accounting for overt subjects displayed in some of them. The internal structure of some infinitival clauses, as well as the contrast between the obligatory subject clitics in some embedded non-finite reflexive contexts and the possible null subjects in non-reflexive entries will lead to a proposal: these embedded infinitives do not have PRO in their subject position.

Finally, Chapter Seven will present some final remarks and a conclusion.
Chapter two

Sentence structure

This extensive chapter deals with decisive questions concerning clause architecture. The final conclusion will be that, for Capeverdean, there are no empirical grounds to propose an immense array of functional categories in the IP domain. On the contrary, every piece of evidence leads to the preferable hypothesis of only one relevant functional head. The label for this head, T, traditionally stands for Tense. According to this more semantically oriented approach, it could stand for Temporal. Either way, it is no more than a label. The relevant point is that there is only one maximal projection in each functional domain.

In order to justify the proposal that the functional domain is this minimal in the language, some core issues on this topic must be addressed.

Namely:

i) the questions about an agreement projection, that is, an AgrP node (this topic is only briefly addressed in section 2.1, since, although an initial observation is needed, it will be developed in greater detail in section 2.4);

ii) the question of whether there is a universal list of functional projections accommodating the particular necessities of all languages, taking the distribution of adverbs (section 2.2) and TMA morphemes (2.3) as diagnostics; this latter section is particularly long, for there are too many details to take into consideration whenever we address temporal reference construction;

iii) the question of whether, on the contrary, we should consider that each language displays only the functional projections needed in order to
account for the respective data (section 2.4); and, in this case, how to accommodate the different markers that help build the temporal-aspectual meaning of the sentence when – as I propose for Capeverdean – there seems to be only a sort of almighty functional projection, TP.

Along the tentative answers to these questions, Capeverdean data will be presented, in order to bring the so-called empirical evidence that sustains the options made. The conclusion, in 2.5, will try to wrap up the complexities addressed in the present chapter.

2.1 Why not an AgrP node

One question we must face when trying to establish a functional structure for a given language is the following: do we need to consider an agreement projection? That is, an AgrP node (or AgrSP, if there is a necessary distinction between subject agreement and object agreement – this last one would be AgrOP). The topics concerning agreement, such as agreement morphology and the abstract Nominative Case, will be considered in greater detail in section 2.4, and also in the next chapters, especially Chapters Five and Six, where null subjects and subject positions, in finite and non-finite clauses, are discussed.

At the present point, though – and before analyzing the Capeverdean temporal reference strategies and the implications they have on the functional architecture of the language –, it is crucial to make it clear that in Capeverdean there seems to be no obvious reason to project an AgrP node.

There are two possible lines of arguments concerning this:

i) the language displays no overt verbal agreement morphology (the verbal paradigm can be illustrated with the verb kanta ‘sing’: N kanta ‘I sang’; bu kanta ‘you sang’; e kanta ‘he/she sang’; nu kanta ‘we sang’; nhos kanta ‘you sang’; es kanta ‘they sang’); this means that there are no strong agreement features for the verb to check through movement to such a functional position;
ii) no available positional arguments favor such a projection, since, as opposed to French, for instance, there is no need for two different heads as landing sites for any verb unit; that is to say that there is no specific word order pattern justifying it.

The debate around the positing of an AgrP for these positional reasons started in Pollock 1989, a seminal paper on verb movement in French and English (lexical finite verbs do not move out of VP in English). The debate is long and complex (including the proposal in Belletti 1990 that the AgrP node dominates the TP node, and not the opposite, as in Pollock 1989), and I shall not go into its details here.

It suffices to say that Capeverdean auxiliary constructions have been analyzed in Baptista 2002 as biclausal structures, a proposal that I have been following in my previous works, and that I will sustain in the present dissertation for some constructions with two verbal entities. This means that in these constructions we have two TPs; the embedded TP is more or less defective depending on the lower (lexical) predicate in question (this will be developed further in the present and also the next chapters). This proposal accounts for the cases in which an auxiliary (like *sta* ‘be’) appears between the negation and the adverb, for instance the temporal *sempri* ‘always’ – notice, however, that this form has been attested by the author, but my informants consider it a borrowing from Portuguese; their much more productive alternatives are *tudu dia* or *tudora*.

Assuming a minimalist approach (Chomsky 1995), one possibly strong feature that is uninterpretable at the interface level is the Nominative Case feature on T (we shall review this notion ahead, in the next chapter). This can be checked (and deleted, and also erased) by attracting the DP-subject from its VP-internal position to Spec,TP (we shall see that, in the presence of *dja* and whenever the subject is a clitic form, it cliticizes to the right of *dja*). There being no other uninterpretable strong features in this domain, no other movement to a functional head in the TP domain is needed.

The relevant point here is that, given the lack of overt verbal agreement morphology, and also in the view of the discussion around positional arguments and syntactic features in the need for being checked, I have found no empirical reason in favor of an AgrP projection in Capeverdean. More evidence supporting this view will be provided in section 2.4, and also in the next chapters.
2.2 Are adverbs reliable diagnostics?

Besides the agreement discussion, the distribution of adverbs has likewise been one of the matters involved in the broad debate regarding the quantity and configuration of functional projections in Universal Grammar. Some proposals (Alexiadou 1997, Cinque 1999, a.o.), more syntactically based, predict a broad range of functional projections, which accommodate any possible functional morpheme and also allow a one-to-one licensing relationship between adverbs and specific functional heads (the adverbs surface on these heads’ specifier positions).

Other authors propose a more semantically-based approach to the distribution of adverbs (Ernst 2002, Costa 2004, a.o.), defending a scope theory against a feature theory and predicting that adverbials appear in adjunction, either to the left or to the right of different positions. Several syntactic rules and constraints are involved (Directionality Principles and Weight theory), but the main point is that the selectional (lexical) properties of adverbs play a crucial role in their distribution. As for markers for Tense, Modality and Aspect, some authors (Bobaljik 1995, Bobaljik & Thrainsson 1998, a.o.) propose a Split, IP parameter, predicting that languages may have only the functional projections needed in order to accommodate their functional morphemes and check their relevant features.

In the present section, I will summarize the main arguments for the two types of approaches concerning the position of adverbs: the one that predicts a universal array of functional projections versus the one that defends a more economical functional architecture (minimalist goals are more respected by the latter). In subsection 2.2.1, I will show Capeverdean restrictions on taking adverbs as diagnostics for a rich functional domain; in subsection 2.2.2 the discussion is centered on the hypothesis of adverbs with an identical meaning behaving uniformly – this hypothesis is not supported by European Portuguese (as in Costa 2004) nor Capeverdean (as in Costa & Pratas 2005).
2.2.1 Capeverdean restrictions

A short summary of the reasons why Ernst 2002 and Costa 2004 approaches to adverb distribution apply to Capeverdean, while Cinque 1999 predictions do not hold, is presented in this subsection.

In Cinque 1999’s system, the main arguments for taking adverbs as diagnostics for a rich universal functional domain are the following:

i) the interpretation of adverbs depends on their position;
ii) there is a (universal) hierarchy for adverb placement;
iii) this hierarchy correlates with similar hierarchies found for heads.

We can follow the details of this hierarchy in (7).

(7) Cinque’s adverb hierarchy:

\[
\begin{array}{c}
\text{frankly} \quad \text{Mood}\text{predicative} \\
\text{fortunately} \quad \text{Mood}\text{evaluative} \\
\text{allegedly} \quad \text{Mood}\text{evaluative} \\
\text{probably} \quad \text{Mod}\text{semantic} \\
\text{once} \quad \text{T(Past)} \\
\text{then} \quad \text{T(Future)} \\
\text{perhaps} \quad \text{Mod}\text{semantic} \\
\text{necessarily} \quad \text{Mod}\text{semantic} \\
\text{possibly} \quad \text{Mod}\text{semantic} \\
\text{usually} \quad \text{Asp}\text{semantic} \\
\text{again} \quad \text{Asp}\text{semantic} \\
\text{intentionally} \quad \text{Mod}\text{semantic} \\
\text{quickly} \quad \text{Asp}\text{semantic} \\
\text{already} \quad \text{T(Anterior)} \\
\text{no longer} \quad \text{Asp}\text{semantic} \\
\text{still} \quad \text{Asp}\text{semantic} \\
\text{always} \quad \text{Asp}\text{semantic} \\
\text{just} \quad \text{Asp}\text{semantic} \\
\text{soon} \quad \text{Asp}\text{semantic} \\
\text{characteristically?} \quad \text{Asp}\text{semantic} \\
\text{well} \quad \text{Voice} \\
\text{fast/early} \quad \text{Asp}\text{semantic} \\
\text{again} \quad \text{Asp}\text{semantic} \\
\text{often} \quad \text{Asp}\text{semantic} \\
\text{completely} \quad \text{Asp}\text{semantic} \\
\text{tutto} \quad \text{Asp}\text{semantic} \\
\text{completely} \quad \text{Asp}\text{semantic} \\
\text{Aspgeneric/progressive} \\
\text{almost} \quad \text{Asp}\text{semantic} \\
\text{completely} \quad \text{Asp}\text{semantic} \\
\text{tutto} \quad \text{Asp}\text{semantic} \\
\text{completely} \quad \text{Asp}\text{semantic} \\
\end{array}
\]

(Cinque 1999:106)

Contrary to this view, there is the one maintaining that adverbs are generated in adjunction configurations, and that their distribution is partly constrained by their lexical semantics, by their prosodic weight and/or by general syntax-semantics interface conditions (Costa 2004, Ernst 1985, 2002).

As described in Costa & Pratas 2005, one of the problems of the specifier approach for adverbs is the following: if adverbs are specifiers of designated functional heads, their position is supposed to be universal. In Capeverdean, however, adverbs do not occur in sentence-medial positions:

(8) a. * Djon ta bai tudu dia skola. (temporal, between verb-locative)
   Djon TMA go every day school
   ‘Djon goes to school every day’.

---

8 For reasons of space and time, I will present only the relevant aspects – and not a detailed description – of these authors’ proposals.
b. * Djon *tudu dia* ta bai skola. (temporal, between subject-TMA)
c. Djon ta bai skola *tudu dia*.
d. *Tudu dia* Djon ta bai skola.

(9) a. *Djon sta* _oxi_ *duenti_. (temporal, in predicative construction)
   Djon be today sick
   ‘Today Juau is sick.’
b. ***Djon _oxi_ sta duenti.***

c. ***Djon sta duenti _oxi_.***
d. _Oxi_ Djon sta duenti.

(10) a. *Nha _fidju_ *so* *gosta di karne.* (object-orient., between subj-verb)
   1SG.POSS son only like of meat
   ‘My son only likes meat.’
b. ***Nha _fidju_ gosta _so_ di karne.***

c. E _so_ di karne _ki_ nha _fidju_ gosta. (object)
   Is only of meat REL my son like

d. (Na nha kasa) E _so_ nha _fidju_ _ki_ gosta di karne. (subject)
   (At my place) Is only my son REL like of meat

(11) a. *Maria ta *papia _dretu_ portuges.* (manner, between verb-object)
   Maria TMA speak right Portuguese
   ‘Maria can speak Portuguese’

This prohibition of adverbs in sentence-medial positions shows not only that they do not surface on any specifier of functional heads (otherwise it would be hard to explain how they never appear between the subject and the verb), but also that one of the possible

---

9 This is possible with a focalized reading of _oxi_ like, say, ‘Today (that I need him so much) John is sick.’

10 This is one of the few contexts in which a weak adverb, _so_ ‘only’, is somehow allowed between Verb and Object, but this is refused by most of the informants from Santiago Island. The most common is for them to present a paraphrase to this, like _Nha _fidju_ e _tudu_ pa karne_ ‘My son is all for meat.’
adjunct positions (following the alternative proposal in Ernst 1985, 2002 and Costa 2004) must be to the right of VP (as is the case for manner adverbs, like in (11b)).

These Capeverdean contrasts render it difficult to maintain that adverbs surface universally as specifiers in the functional domain. In the next subsection, a more flexible approach, in the fashion of Ernst 1985, 2002 and Costa 2004, is presented.¹¹

2.2.2 Meaning, prosodic weight, lexical specifications and categorial status

In specific cases of Capeverdean, and given some properties of prosodic weight (and possibly of categorial status as well), certain temporal adverbs such as oxi ‘today’ are preferred to the left of TP – the examples in (8c) and (9c) display different constraints for two adverbs semantically related. The example in (9c) would be okay if oxi were focalized, like in: Juau sta duenti oxi, mas manha e ta ben trabadja. ‘Juau is sick today but tomorrow he comes to work.’

On the other hand, the analysis in Costa 2004 shows that in European Portuguese adverbs with related meanings may distribute differently depending on their lexical specification. In particular, adverbs which are ambiguous between a subject-oriented reading and a manner reading exhibit a more restricted distribution than those that are not ambiguous. This fact makes it possible to argue that adverb placement and interpretation are partly constrained by lexical properties. This view favors Ernst’s 2002 proposals as well.

There is a classical type of counter-argument for symmetric adjunction that makes use of the following contrast:

(12)  a. John stupidly spilled the coffee. Subj-Or/*Manner
     b. John spilled the coffee stupidly. *Subj-Or/Manner

¹¹ See Lobo 2003 for a detailed description on the position of adverbial clauses in European Portuguese, in relation with its informational status; the author points out that ‘background’ clauses are adjoined to the left (of TP, for instance), on a peripheral position, and that clauses containing information which is not expected (which, thus, can be focalized), are adjoined to the right of a lower position, for instance VP (they are internal to TP). Hence, the relevant feature is [presupposition].
Ernst (1985, 2002) proposes that syntax generates symmetric adjunction, but some configurations may be filtered out by semantic constraints, in particular when there is lexical ambiguity. As said before, in the same line of reasoning Costa 2004 argues that when there is no lexical ambiguity, symmetry is not problematic. Hence, we have the contrast between the European Portuguese sentences in (13), where there is lexical ambiguity, and in (14), when there is no lexical ambiguity, showing that the interpretation of the adverb is not primarily conditioned by syntax:

(13) Subject-oriented reading (ambiguous adverb):
   a. O João estupidamente tinha estado a falar com os amigos.
      the João stupidly had been talking to the friends
   b. ??O João tinha estupidamente estado a falar com os amigos.
      the João had stupidly been talking to the friends
   c. *O João tinha estado estupidamente a falar com os amigos.
      the João had been stupidly talking to the friends
   d. *O João tinha estado a falar estupidamente com os amigos.
      the João had been talking stupidly to the friends
   e. *O João tinha estado a falar com os amigos estupidamente.
      the João had been talking to the friends stupidly

(14) Subject-oriented reading (non-ambiguous adverb):
   a. O João propositadamente tinha estado a falar com os amigos.
      the João intentionally had been talking to the friends
   b. O João tinha propositadamente estado a falar com os amigos.
      the João had intentionally been talking to the friends
   c. O João tinha estado propositadamente a falar com os amigos.
      the João had been intentionally talking to the friends
   d. O João tinha estado a falar propositadamente com os amigos.
      the João had been talking intentionally to the friends
   e. O João tinha estado a falar com os amigos propositadamente.
      the João had been talking to the friends intentionally

These data in Costa 2004 bring additional support to the minimalist consideration – also crucial at various points of the present work – that syntax generates multiple outputs,
which may be filtered at the interface with interpretation, in order to respect lexical constraints defined independently of the syntactic component.

Another reason against the universality proposal is that in European Portuguese certain adverbs display a head-behavior, distributing differently from their XP counterparts.\(^{12}\) This fact argues in favor of the view that, besides prosodic weight, and also besides lexical specifications such as directionality and ambiguity, also the categorial status is relevant in some languages to predict different adverb distributions for the same semantic class of adverbials.\(^{13}\) The line of reasoning goes as follows: the position of certain adverbs does not depend on their meaning, but rather on their prosodic weight and/or their categorial status. Notice that there is no evidence to say that this position is derived via movement.

Finally, we must underline that any intervention of any adverb in between Capeverdean TMA markers, or between TMA markers and the verb, is disallowed.

(15) a. *Djon **ta** oxi kume na kuzinha. (temp. between TMA and V)

Djon TMA today eat in kitchen

‘Today Djon eats in the kitchen.’

b. Oxi Djon **ta** kume na kuzinha.

From (15a) we may infer that there are adjacency requirements hard to explain under an exploded functional domain. If some given specifiers of functional heads were available for an adverb like oxi (a relatively prosodically light expression, compared with tudu dia, for instance – this means that there are no reasons of “excessive” weight involved here), it would be difficult to explain the prohibition in (15a).

\(^{12}\) As said before, the examples that illustrate the possible sentence-medial position of adverbs are rare in Capeverdean, with the possible so ‘only’ between the Verb and an Object PP (10b) being one true exception. In any case, so does not even have a semantic counterpart with a different categorial status, in order to compare their different position constraints. Following some different data in Baptista 2002, which include the adverb ben as a counterpart to dretu, let us consider a counterpart to the example in (11a) with ben to the left of the object (Maria ta papia ben portuges). The pair ben/dretu could, then, be an example of this distinction between head adverbs and their XP counterparts. The version ben for the manner adverb dretu, however, is systematically rejected by my informants, who recognize it as Portuguese.

\(^{13}\) For a discussion on this topic for European Portuguese adverbs, see Castro & Costa 2003.
I hope to have shown that the Capeverdean data involving adverbs make it difficult to maintain that these behave uniformly cross-linguistically, as it is predicted by the universal functional hierarchy analysis of Cinque 1999’s system. Instead, these data follow more straightforwardly from more flexible approaches to the syntax of adverbs (Ernst 1985, 2002; Costa 2004).

The main point is that any significant information on the functional architecture of a language must be provided by some diagnostics other than adverb placement, since the results that this one provides are not reliable. This is one of the points I will try to capture in the next section. The proposal will be that there is no reason to consider any maximal projection other than TP; this view is supported by a more reliable diagnostics: the implications of the temporal reference construction strategies in the language.

2.3 TMAs and other functional units

In section 2.1 I have stated that there is no evidence in favor of an AgrP node, either for morphological reasons (there is no overt agreement morphology) or for positional reasons (there is no need for an Agr head to accommodate the verb, either auxiliary or not)\(^\text{14}\). Following Ernst 2002 and Costa 2004, in 2.2 I have discarded adverbs as a reliable diagnostics with respect to the functional structure.

In the present section I will follow a different possible line of inquiry regarding Capeverdean functional architecture. It relates to the organization of the functional morphemes, the so called TMAs, that combine (both with each other and with other operators, such as adverbials and other expressions, as well as with the inherent temporal properties of the predicates themselves) to build temporal reference, and also other units that surface in the functional domain, such as the negation marker \(ka\), the copula \(e/era\), and the subject clitics. In spite of having no independent semantic content – this is by definition a property of functional words –, these functional morphemes bring relevant information to the sentence: temporal information in the case of TMAs; negation in the case of the morpheme \(ka\) (which, as we shall see, is likewise implicated in some circumstances of temporal marking) and, in the specific case of subject clitics, some other encoded features and semantic information, such as Case and theta-roles.

\(^{14}\) We shall see that Baptista 2002 argues that an Agr head is needed for a different reason: it is where subject clitics surface, licensing a \(pro\) in the specifier position. I will argue, in subsection 2.4.1, that this seems a resort in order to license \(pro\), but in face of the fact that there is no other empirical evidence to consider there is a referential \(pro\) in the language, again an Agr head is not needed.
Consider the contributions of \textit{ta} and \textit{-ba}, the two temporal morphemes that seem most challenging with respect to the diversity of temporal environments they enter into.

From (16), one could infer that preverbal \textit{ta} means Present (tense) Habitual (aspect).

(16) \textit{Txoti ta ben kume na jardin.} 

\begin{itemize}
  \item birds TMA come eat in garden
  \item ‘The birds come in the garden to eat.’
\end{itemize}

In (17), however, it looks more like a Future:

(17) \textit{Oxi N ka leba nha fidju trabadju, mas otu dia N ta leba-l bombudu.} 

\begin{itemize}
  \item today 1SG NEG take my son work but anotherday 1SG TMA carry.3SG on back
  \item ‘Today I didn’t take my son to work, but one of these days I will carry him on my back.’
\end{itemize}

The absence of \textit{ta} in the first clause results in the Past reading for \textit{leba} ‘take’; \textit{ta} in the second clause, together with the expression \textit{otu dia} ‘another day’, has a Future reading; we can confirm that, as it seems now, this morpheme means non-Past.

With some other verbs, such as \textit{sabe} ‘know’, the lack of \textit{ta} does not result in a Past reading:

(18) \textit{N ka sabe p-undi ki e sta.} 

\begin{itemize}
  \item 1SG NEG know to-where REL 3SG BE
  \item ‘I don’t know where he is.’ (in the sense of these days, not of this very moment)
\end{itemize}

With \textit{sabe} ‘know’, \textit{ta} introduces a different value from the one obtained in (16):

(19) \textit{Si bu mexe na kes livru-li, bu pai ta sabe.} 

\begin{itemize}
  \item if 2SG touch in DEM books-LOC POSS.2SG father TMA know
  \item ‘If you touch these books, your father is going to know.’
\end{itemize}
As for the postverbal -ba, with sabe ‘know’ and with sta ‘be’, one could infer that it marks Past:

(20) (Onti) N ka sabe ba undi ki bu sta ba.
yesterday 1SG NEG know.TMA where REL 2SG BE.TMA
‘Yesterday I didn’t know where you were.’

Nevertheless, with verbs like kore ‘run’ or fla ‘say/tell’, there are different implications for -ba. It combines with ta for a Past Habitual reading (21a). It occurs without ta in embedded clauses which seem non-finite (21b), but it is barred in these same non-finite embedded contexts when some other element in the sentence has been changed, namely the temporal reference of the matrix clause: the matrix verb in (21c) is marked for some tense like Past Anterior. We might think that the negation could have some influence in this -ba selection, but from (21d) we may conclude that this is not the case:

(21) a. Kantu Pedru era pikinoti, e *(ta) kore ba txeu.
   when Pedru BE.PAST little.kid 3SG TMA run.TMA a.lot
   ‘When Pedru was a little kid, he used to run a lot.’

b. Si mai ta fla ba el p’-e kore ba.
   POSS3SG mother TMA tell .TMA 3SG PREP-3SG run.TMA
   ‘His mother used to tell him to run.’

   but 2SG no 1SG tell .TMA 2SG PREP-2SG NEG run.TMA
   ‘But not you. I had told you not to run.’

d. Otu algen ta fikaba la n-igreja pa ka more ba.
   other people TMA STAY.TMA LOC PREP-church PREP NEG die.TMA
   ‘[when there were floods] other people would take refuge in the church, so that they wouldn’t die.’

In some embedded contexts which look like non-finite environments, ta is also needed, yielding some value of Progressive plus simultaneity (with respect to the reference time – the one introduced in the matrix clause).
(22)  
\textit{Bu flaba mi ma bu ka ta fumaba, mas N odja-u ta fuma.}

\hspace{1em}2SG say.TMA 1SG COMP 2SG NEG TMA smoke.TMA but1SG see-2SG TMA smoke

‘You had said you didn’t smoke, but I saw you smoking.’

Other Progressive contexts are built with \textit{sata} (either in Present (23a) or Past (23b), the latter being also marked with \textit{-ba}). A Future reading (23c) is also available for \textit{sata}.

(23)  
\begin{enumerate}
  \item \textit{Miriam, bu sata kume bolu antis di djanta?}
  \hspace{1em}Miriam, 2SG TMA eat cake before of dinner
  \hspace{1em}‘Miriam, are you eating cake before dinner?’
  \item \textit{Kantu-N txiga, es sata komesaba ta kanta.}
  \hspace{1em}when-1SG arrive 3PL TMA start.TMA TMA sing
  \hspace{1em}‘When I arrived they were starting to sing.’
  \item \textit{Sata txobe oxi.}
  \hspace{1em}TMA rain today
  \hspace{1em}‘It’s raining today. / It is certainly going to rain today.’
\end{enumerate}

In the first two subsections of this section, it is my goal to show that Capeverdean challenges any proposal of establishing a hierarchy of functional projections for TMA morphemes. The first – and main – reason for this is that, as we have just seen in this short sample of sentences, the same functional morpheme can bring different meanings to slightly distinct environments. Each of these semantic inputs of a given morpheme depends on what is happening in the rest of the structure.

Hence, all we can hitherto know for sure about TMAs is that:

\begin{enumerate}
  \item there is a limited set of functional morphemes in the language (preverbal \textit{ta} and \textit{sata}, besides the more particular \textit{dja}; postverbal \textit{-ba});
  \item each of them may appear by itself or coexist with (some) others;
  \item when there is more than one morpheme, they appear in a certain order and obey some given restrictions (such as: only one of them, \textit{-ba}, appears post-verbally);
  \item the contribution that each brings to the sentence depends on the semantic properties of predicates and on the presence/absence and meaning of other functional morphemes, adverbials and discourse information.
\end{enumerate}
These facts condemn to failure any tentative of a well behaved hierarchical structure with each of these morphemes surfacing under a distinct label.

In other words, and given the language strategies for the construction of temporal reference (here I am including in ‘temporal’ also the aspectual and modal reference), it would be at least dangerous to propose a direct mapping of each functional unit that participates in this reference construction onto particular functional projections such as AuxP, AspP, MoodP, among others.

More specifically, in most Capeverdean sentences Tense and Aspect, and also Mood, are not exclusively provided by functional morphemes, but rather they are derived from the interaction of different pieces, such as the verbs that these TMAs are marking, adverbial expressions, temporal clauses and discourse information, which work together and condition the whole meaning.

As for the Mood, there are two important reasons for me not to discuss this topic in detail, at least in the present work. One of them is that in Capeverdean any remark on the Mood of some verbal construction has complex implications that must be subject to further accurate research. There is no specific verbal morphology that identifies subjunctive, although there are sentences whose reading clearly is correspondent to this mood in Portuguese – in Portuguese it is identified by specific morphological markers. We shall see that in some counterfactual environments, for instance, the two morphemes involved are preverbal \( ta \) and postverbal \(-ba\), exactly the same that participate in past habituals, either in a durative or an iterative reading. As for conditional, there is a wide debate around this being really a mood; hence I will not even consider this possibility. This does not mean, though, that there are no conditional constructions – they exist in Capeverdean, as we shall see, but the morphology involved is the one available for indicative mood; as it might be expected by now, the conditional reading is provided by other elements in the sentence (such as a conditional clause introduced by \( si \) ‘if’).

Modal reference – like non-factuality, encoded in the subjunctive in some languages – is sometimes built in Capeverdean with recourse to modal auxiliaries (such as \( pode \) ‘can’, and these cases will be identified locally), but also to the combination of certain morphemes, as just mentioned, together with pragmatic and discourse information.

Aspect and Tense, besides being dependent on the interaction of these pieces of information, are widely considered as being temporal, whereas modality is viewed as
being something different: a category of linguistic meaning related to the expression of possibility and necessity (von Fintel 2005, to appear). And this – the fact that it is somewhat distinct from the temporal reference construction – is the second reason for me not to discuss this topic in detail here (interspersed references on this will be included when needed).

This means that, when discussing the mapping of functional morphemes onto syntax, and the implications of this in the clause architecture of Capeverdean, I will refer to each of these morphological pieces as TMA as a notation strategy (which will leave the modality discussion open at this point), focusing mainly on the temporal reference construction; that is, Tense and Aspect.

All this said, I shall argue that syntactic Tense features and also morphological markers for Tense and Aspect (TMAs) may be accommodated by T alone: besides describing a wide range of new data (which will also be crucial for some other topics in the next chapters), the main purpose in the present chapter is to show that each of these markers may surface in one of the multiple adjoined heads under the label T. Their position relatively to each other must be accounted for by requirements other than hierarchic ones; these other requirements may be connected with linearization and semantic relations (namely scope)\(^{15}\).

This proposal is grounded on the evidence that, in Capeverdean, temporal semantics is, as argued earlier (although in different words), typically compositional\(^{16}\). In order to deal with this compositionality, in subsection 2.3.1 I will summarize some relevant theoretical tools on predicates and their relation to time perspective, that is, Aspect. In subsection 2.3.2 I will apply these theoretical observations to Capeverdean data. As for Negation, in subsection 2.3.3 I will try to show how there are empirical motivations for contending it surfaces on a T head as well: it switches positions with the copula according to the Tense that this one encodes (the negation marker *ka* appears post-copula whenever the copula is the present form *e ‘is’, and pre-copula whenever

\(^{15}\) This in turn may be related to more complex facts, such as the availability of only some of these morphemes in certain embedded contexts, as we shall see further.

\(^{16}\) In Baptista 2002 there is a contrast between, on one hand, the detailed description of data (Baptista 2002, chapter 4), accounting for the complex interdependencies of Capeverdean verbs and the different TMAs, and, on the other hand, the proposal, in chapters 6 and 7, of a Split, IP structure, where, for instance, *ta* surfaces under an Asp label on page 166 and under an Aux label on page 200, for exactly the same sentence: *Joao ta staba ta kumeba ‘Joao would have been eating’* (author’s translation). The present work follows the spirit in the author’s chapter 4 and presents a proposal different from the author’s chapters 6 and 7.
this one is the past form *era* ‘was’/*used to be’); furthermore, only in Present negation contexts can the copula be null. Subject clitics, whose position must be connected with the one of the puzzling marker *dja*, will also be discussed in the point 2.4.2.2, in the next section.

### 2.3.1 The true value of Aspect

Events are located in time. Besides this temporal location, which is encoded in Tense, events are perceived through a time perspective. Aspect is what encodes the type of temporal information related to perspective (events can be perfective/imperfective, durative/instantaneous, among many other properties that may be, or not, organized in pairs). Consider the following sentences:

(24)  
\[ \text{a. Onti N txiga sedu.} \]
\[
\text{yesterday 1SG arrive early} \\
\text{‘Yesterday I arrived early.’} \\
\]
\[ \text{b. Kantu N era nobu, N ta duraba ku txiga tudu dia.} \]
\[
\text{when 1SG was young 1SG TMA take.long.TMA with arrive every day} \\
\text{‘When I was young, I always arrived late.’} \\
\]

Both sentences are located in the Past, defined relatively to the utterance time. (24a), however, is an episodic occurrence, whereas (24b) is an habitual occurrence (not in the durative, but in the iterative sense – which is given by some internal properties of the verb itself, for it is impossible that someone is continuously in the process of arriving in a continuous period of time).

This habitual occurrence, in (24b), besides being located relatively to the utterance time (Past), is also defined relatively to a reference time: the youth of the speaker. The period of time during which the speaker kept arriving late is coincident with (overlaps) the period when he was young (nothing is said on whether the boundaries of both periods are coincident or not).

This last type of temporal information, Aspect, has been extensively studied especially among semanticists, with recent syntax studies also giving great attention to it, more
specifically to a crucial question concerning the syntax/semantics interface: how is Aspect mapped onto a syntactic structure?

Before proceeding with the analysis of Capeverdean empirical evidence – which will be the case in subsection 2.3.2 – some relevant concepts concerning Aspect theories and analyses must be summarized.

This summary – in the present subsection – does not have the intention of being exhaustive regarding all the possible topics. On the contrary, this selection has been guided by the main goal in this chapter, which is to cross different theoretical tools and Capeverdean data that support the following two points:

i) a (derived) compositional approach of Tense and Aspect reference;

ii) only one maximal projection is necessary in the functional domain to accommodate the distinct functional morphemes.

These theoretical concepts will be summarized in two separate, more descriptive points, for reasons of clarity and organization. The main concepts supporting the point i) will be in 2.3.1.1, and the main concepts supporting the point ii) (although, naturally, they have implications also for i) ) will be in 2.3.1.2.

2.3.1.1 Temporal perspective(s): facts for a compositional approach

There are two types of Aspect information (the way events relate to time):

- A. Viewpoint: also called grammatical Aspect, it locates events in a relation with the reference time – events may be Perfective or Imperfective (and the Imperfective may have a Progressive or a Habitual – either durative or iterative – reading)\textsuperscript{17}. Generative Grammar seeks the universal constraints on viewpoint and the properties connected to cross-linguistic variation.

\textsuperscript{17} For discussions on this, see Comry 1976 and subsequent works, Dowty 1979, Hornstein 1990, Smith 1997, a.o.
- B. Aktionsart: also called lexical Aspect, it relates to the internal temporal constituency of events (beginning/ending, involvement or not of a change of state, etc.)

In the Aktionsart field, we can refer to “situation types” / “eventuality types” – these are the more common generic expressions keeping the internal distinction between “non-states” and “states”. Nevertheless, the generic name “events”, as applied to “eventive predicates” (non-states) and also “states”, has become increasingly accepted. An extended classification includes (cf. with Vendler 1957, Smith 1997):

Situations / Eventualities (or, also generically, Events)

i) states (such as sta kontenti ‘be happy’);

ii) eventive predicates:
   - activities (such as kore ‘run’);
   - accomplishments (such as skrebe un karta ‘write a letter’);
   - achievements (such as txiga ‘arrive’);

The division into two groups is due to the following assumptions: “states” refer to qualities (sabe ‘know’, ten/tene ‘possess’/’have’ or, in some cases, gosta ‘like’ something, among many others), whereas the verbs in the second group refer generally to different types of “actions” (except for certain kinds of “achievements”, such as ‘to recognize something’). There are important distinctive properties among the second

---

18 There are views that put the main burden of Aspect on conceptual structures, such as Jackendoff 1996. For views relating traces of Aktionsart and structural properties of predicates, see Borer 1998, 2005, Ramchand 1997, 2003, a.o.

19 For an analysis of states in European Portuguese, see Cunha 1998. Analyzing the possible combinations of progressive constructions with some types of predicates, the author proposes a division within the “states” group: those which can be phased, and thus can behave like non-states, and those which cannot. Confront for instance ‘he is being nice’ with ‘*he is being tall’.

20 There is a distinction in the Capeverdean forms for ‘possess’, ten, and for ‘have’ (momentarily), tene. Thus, the following sentence is possible: Mi e riku mas N ka tene dinheru ‘I am rich but I do not have money [with me/ now].’ The form ten is also the one in use in:
   - existential constructions, such as Ten tres katxor na nha rua. ‘There are three dogs in my street.’
   - deontic modal constructions, such as N ten ki ba kasa sedu. ‘I must go home early.’

21 Also perceptive verbs, like ‘see’ and ‘hear’, depending on the Aspect marking, may be considered states. Notice also that, in fact, the “states” group is not homogeneous, as we shall see for Capeverdean as well (although I will not propose a classification for the different verbs).
group: for instance, we may say that “activities”, such as kore ‘run’, are [-resultative] and “achievements”, such as txiga ‘arrive’, are [+resultative]. But properties like this, and also like the terminative/durative opposition, depend often on the Aspect marking combined with the inherent properties of predicates\(^{22}\).

In other words, there are various common/distinctive levels that one can use to a taxonomic diagnostics. This is a wide field of research which has produced a broad range of different studies and proposals.

The most interesting crossroads is this one between Aktionsart and Viewpoint: the implications and restrictions with respect to the possible tenses and viewpoint properties according to the different groups of predicates. For instance, according to Vendler 1957, whereas “activities” and “accomplishments” are processes going on in time (consider, for instance, *N sata kore* ‘I am running’, *N sata skrebe karta* ‘I am writing a letter’), “states” and “achievements” are not processes, hence they can never be applied continuous tenses (**‘I am knowing’, **‘I am believing’, etc.)\(^{23}\); “states”, as qualities, can be predicated for shorter or longer periods of time, whereas “achievements” are connected with single moments (although anda ‘walk’ or kore ‘run’ may have been processes leading to the ‘arrival’, there is only one instant which we can refer to as the ‘arriving’ itself).

Besides the interaction between Aktionsart and some TMA morphology, there is the interaction between both these Aspect types of information and certain prepositional adjuncts – this is illustrated in English by the examples in (28), ahead.

Now that some basic concepts and ideas have been introduced, a few more distinctions with respect to the temporal constituency of events are relevant for the present work. Among other classifications, events may be labeled according to the following pairs:

---

\(^{22}\) Cunha 1998 points out that, for instance, typical “achievements” in Portuguese, such as *ganhar a corrida* ‘win the race’, may acquire a durative property (hence, not terminative, as is typical from “achievements”) when combined with certain temporal references, namely progressive: *O João esteve a ganhar a corrida durante uma hora* (mas não a ganhou) ‘John has been winning the race for an hour (but he did not win it).’

\(^{23}\) There are, however, possible combinations of progressive with some “states”, such as the proposed in Cunha 1998 (see fn. 19, previous page), concerning stative constructions with ‘be’. We shall see that Capeverdean *gosta* ‘like’, also considered a stative verb, may combine with progressive morphology as well, as opposed to *sabe* ‘know’. 
i.) static/dynamic;

ii.) telic/atelic (telicity: a change of state is associated with the end of the event);

iii.) durative/instantaneous.

These classifications often meet many difficulties in the real life of predicates and events. Most times we need to essay an approach that considers, for instance, the common features between accomplishments, such as *kore un kilomitru* ‘run a kilometer’, (which involve a process component and a change of state) and, on one side, activities, such as *kore* ‘run’ (process component) and, on the other side, achievements, such as *txiga* ‘arrive’ (a change of state). We need to consider as well that in some languages the addition of some prefixes or direct objects induces telicity to an otherwise atelic predicate. This implies that a compositional approach is more adequate.

According to Verkuyl 1993 and Krifka 1998 (the important differences between these approaches are not relevant here), only to cite two studies on this particular topic, there are some specific facts motivating a compositional approach to Aktionsart. One of these facts relates to the effects triggered by the nature of some constituents, such as direct objects and prepositional adjuncts.

Some direct objects have this effect when the event descriptions involve predicates with an “affected theme”. One such predicate is ‘draw’: if the affected theme/direct object is a count expression, the event is an accomplishment; if it is not, the event is an activity. We can see this in the contrast:

(25) a. ‘Peter draw two clouds.’ (accomplishment)
    b. ‘Peter draw clouds.’ (activity)

Thus, according to Verkuyl 1993, some verbs have a thematic relation with their objects such that the object is totally affected. In other words, the terminative / durative property of the event description depends on the count expression / non-count expression nature of the object.

The event description for other predicates, which do not have an affected theme (their thematic relation with their objects determines that the object is partially affected), is not this dependent on the nature of direct objects: *pintxa kel karinhu li* ‘pushed this cart’ / *pintxa karinhus* ‘pushed carts’ (both activities).
In these descriptions, one crucial notion is the quantization of predicates: a predicate P is quantized if whenever it applies to an object, it does not apply to a proper subpart of the object. (Krifka 1998:200)

Quantized predicates correspond to telic predicates, and involve countable individuals – one possible example of the countable/non-countable opposition: cakes/wine. Another way of saying this: a telic predicate has quantized reference (26a); an atelic predicate has a cumulative reference (26b). As for (26c), it is ambiguous, since we may understand that Pedru drank all the wine he had in his glass (telic), but also that he drank from his wine, and not anyone else’s (atelic).

(26) a. Pedru kome tres bolu.
   ‘Pedru ate three cakes.’

   b. Pedru bebe vinhu na djanta.
   ‘Pedru drank wine at dinner.’

   c. Pedru bebe si vinhu [dumbes].
   ‘Pedru drank his wine [at once].’

One more relevant notion at this point is the subinterval property of predicates: a predicate P has the subinterval property if whenever P holds of an event/interval, it also holds of every sub-event/sub-interval of that event/interval. States and activities have the subinterval property. Let us check the following pair:

(27) a. Pedru kore na jardin. (predicate P)
   ‘Pedru run in the garden.’

   b. Pedru kore pa jardin. (not predicate P)
   ‘Pedru run to the garden.’

In (27a) we have an atelic property; in (27b) we have a telic property (a change of state associated to the end of the process: Pedru arrived to the garden).

With respect to predicate taxonomy diagnostics related to prepositional adjuncts, we have to consider the contrast between adverbial expressions of the ‘for ten minutes’ type (permitted by atelic predicates and prohibited by perfective telic predicates) and the ‘in
ten minutes’ type (the relation is reversed: these are permitted by some perfective telic predicates and prohibited by atelic predicates). Let us see the examples in (28):

(28) a. ‘Peter has been writing for ten minutes’
    b. ‘Peter has been writing a letter for ten minutes’
    c. * ‘Peter ate two apples for ten minutes’
    d. ‘Peter ate two apples in ten minutes’
    e. * ‘Peter has been writing in ten minutes’
    f. * ‘Peter has been writing a letter in ten minutes’

The ‘for ten minutes’ type of adverbials may occur with predicates that denote processes, like activities (28a) and some cases of what are taken as accomplishments but, by way of certain aspectual marking, may behave like activities (28b). The ‘for ten minutes’ type of adverbials cannot occur with predicates that are marked [-process], like the terminative one in (28c).

Adverbial expressions of the ‘in ten minutes’ type may combine with [-process] predicates, like the one in (28d), but not with activities (either the typical activity in (28e) or the accomplishment rendered activity by way of aspectual marking, in (28f)). Each of these types of expressions has different implications with respect to events.

Notice, also, that there is an ongoing debate on the multiple relations between the underlying value of predicates concerning telicity and the possible combinations through which a different value can be derived.

These and other diagnostics that make use of different types of expressions and sentence constituents in order to classify predicates according to their aspectual properties have problems of their own, since they make the right predictions for some predicates/sentences/languages and wrong predictions for others.

Let us consider other relevant relations that must be discussed when we cross all these notions of Aktionsart with grammatical Aspect (Viewpoint). For instance, the relevant (for the present purposes) association of Perfective with telic and of Imperfective with atelic predicates. The proposal in Swift & Bonhemeyer 2002 considers a default Aspect, which is marked by a zero morpheme in certain environments and, thus, the relevant information must be encoded in the internal properties of the predicate.
This involves the following prediction:

- if a language has verbal forms with no morphological marking for Aspect, and the aspectual meaning of these forms varies according to the telic property of the predicate, therefore:

  i) the zero form of atelic predicates (including activities, since they do not involve a change of state) cannot have a perfective meaning;

  ii) the zero form of telic predicates (which include a change of state) cannot have an imperfective meaning.

As we shall see in greater detail in the next section, an adaptation of this is crucial for Capeverdean. It can be restated as follows:

In Capeverdean:

i) the zero form of most among the so-called stative verbs cannot have a perfective meaning;

ii) the zero form of all predicates that do not denote states (including activities) cannot have an imperfective meaning;

iii) a few verbs are ambiguous; this is the case of *gosta* ‘like’, which sometimes obeys i) (whenever we refer to some personal property, such as *N gosta di pasia parmanhan* ‘I like to take a walk in the morning’), and other times obeys ii) (whenever an adverbial like onti ‘yesterday’, for instance, is involved, and the whole sentence has a clear perfective reading, such as *N ka gosta di filmi onti* ‘Yesterday I didn’t like the movie’).

A different topic much discussed and developed in the last decades, concerning the possible relations between Viewpoint and Aktionsart, is the puzzling English Perfect.

Among the main theories on this we have:

i) the Result State theory\(^\text{24}\) (Perfect means the result state of the event at reference time – sentences containing a Perfect are stative sentences, whereas sentences containing a past perfective may be statives or non-statives, depending on the predicate in question);

\(^{24}\) Parsons 1990, Giorgi & Pianesi 1998 (analysis of the Portuguese Aspect system), a.o.
ii) the Anteriority Theory \(^{25}\) (which locates the event time, \(E\), prior to the reference time, \(R\) – as opposed to the Simple Past, which expresses anteriority of the reference time \(R\) with respect to the Speech time, \(S\));

iii) the Extended Now theory\(^{26}\) (for which the Perfect introduces the Perfect Time Span, whose right boundary is set by Tense and whose left boundary may be set by adverbials).

These proposals have distinct consequences on the syntactic structure organization of distinct elements that build the temporal reference. I mention them here to illustrate the broadness of issues concerning the crossings of Viewpoint and Aktionsart.

More central to the present work is the debate on Imperfective. One essential point to consider cross-linguistically is the possible relation between Imperfective morphology\(^{27}\) and Progressive or Habitual. In some languages, the same form, Imperfective, marks both Progressive (event-in-progress/ongoing event) and Habitual\(^{28}\). In other languages, Progressive and Habitual are built with recourse to distinct strategies. This is the case in Portuguese and also in Capeverdean\(^{29}\). In (29) we have illustrating examples:

---


\(^{26}\) Dowty 1979, Iatridou, Anagnostopoulou & Izvorski 2001, a.o.

\(^{27}\) Present is also non-perfective, although whenever we refer to imperfective morphology we are often referring to past imperfective.

\(^{28}\) This is the case in Spanish, and also in Modern Greek. An example from the latter is the following (from Bhatt & Pancheva 2005):

\[\begin{array}{ll}
\text{Egtiza} & \text{ena} \quad \text{spiti.} \\
\text{build.PAST.IMP.1SG} & \text{a house} \\
\text{‘I was building a house.’ / ‘I used to build a house.’} \\
\end{array}\]

\(^{29}\) This is the case also in Hindi (Bhatt 1999), where there is an “imperfective” morpheme intervening in habituals and a periphrastic auxiliary (\(\text{rah} \ ‘\text{stay}’\)) in progressives. This is illustrated in (i) and (ii):

(i) habitual

\[\begin{array}{llll}
\text{Yusuf} & \text{skuul} & \text{jaa-taa} & \text{hai.} \\
\text{Yusuf school} & \text{go-IMP/HAB} & \text{be.PRS.SG} \\
\text{‘Yusuf goes to school.’} \\
\end{array}\]

(ii) progressive

\[\begin{array}{llll}
\text{Yusuf} & \text{skuul} & \text{jaa raha} & \text{hai.} \\
\text{Yusuf school} & \text{go Prog} & \text{be.Prs.Sg} \\
\text{‘Yusuf is going to school.’} \\
\end{array}\]
(29)  a. Quando era pequeno eu comia muitas bananas.
   a’. Kantu N era pikinoti N ta kumeba banana txeu.\textsuperscript{30}
   ‘When I was a little kid I used to eat lots of bananas.’

   b. Quando chegaste eu estava a comer uma banana.
   b’. Kantu bu txiga N sata kumeba un banana.
   ‘When you arrived I was eating a banana.’

This said, let me summarize the relevant possible crossings between what we know about Imperfective and some notions of Aktionsart:

i) Imperfective (Habitual) morphology is typically able to appear with stative predicates, whereas Progressive morphology (of the periphrastic type, like in Hindi, for instance) typically is not;

ii) Imperfective (both Habitual and Progressive) event descriptions are non-quantized (that is, if they hold for an interval, they hold for subintervals of that interval as well); Perfective event descriptions may be quantized or not, depending on the predicate and the other sentence elements involved.

Whereas the point ii) is easily acceptable from what we know so far about these categories, we shall see that the point in i) needs some interesting adjustments for Capeverdean analyses.

In other words, and only as an introduction, Progressive morphology (sata or, in the periphrastic version, sta ta) is indeed not able to appear with some stative predicates (like sabe ‘know’ and kre ‘want’, which are more typically stative for that matter than gosta ‘like’), either in Present or Past. As for the (overt) Past Habitual morphology (ta

\textsuperscript{30} One peripheral note on txeu, which means ‘a lot’: txeu may appear between the verb and the noun only with countable objects, and even in this case, as in this example with banana, it is sometimes preferred in sentence-final position. Either way, it is txeu that dispenses with the marking for plural on the countable nouns – we shall address briefly the plural marking strategies in Chapter Three). With cumulative objects it certainly appears in the end of the clause, such as in N teneba dinheru txeu ‘I had lots of money.’ One hypothesis is that in the former combination txeu functions like an adjective, inside the object DP, and in the latter it is an adverb, in which case its final position is expected.
V-\textit{ba}), it actually occurs with statives in a particular type of environments: the consequent clauses in some counterfactuals.\textsuperscript{31}

Such as in (30):

(30) \textit{Si bu konkoba (na) porta N ta sabe\textsubscript{a} ma-u staba li.}

‘If you knocked on the door I would know that you were here.’

The morpheme \textit{ta} is able to appear with statives all by itself as well, in Future readings:

(31) \textit{Manhan nu ta sabe si e Pedru Pires ki ganha.}

‘Tomorrow we will know whether it is Pedru Pires who has won [the elections].’

Resuming the summary of the crucial concepts involved in the present work, there are different proposals for the Progressive as well, some of which consider that it should be analyzed in terms of modal semantics. For instance, the assumption in Dowty 1979 that Progressive is a combined modal/temporal operator. The basic intuition here is that a sentence like \textit{Maria sata leba livru} ‘Maria was reading a book’ is true if and only if something was going on that, if it had proceeded normally, there would be a certain point in time where Maria would have read the book. This allowed us to make reference to possibly non-actual complete events.

Parsons 1990 presents an analysis of Progressive that does not involve modality, and introduces incomplete events as primitives. We have the “hold” notion (roughly,

\textsuperscript{31} The classification counterfactual is used here in the sense assumed in Iatridou 2000: it concerns situations that cannot be helped anymore. This is the case of past counterfactuals, usually associated with Pluperfect morphology in languages where it is available. We shall see that in Capeverdean when the affix \textit{-ba} occurs without being combined with preverbal \textit{ta} for non-stative verbs a Pluperfect reading or the equivalent to a subjunctive reading may obtain, depending on the context.

This use of the term counterfactuality accounts for two environments: “counterfactual wishes, whereby the subject expresses a desire for things to be different from what they are or were” (Iatridou 2000:231)

(i) a. I wish I had a car. (conveys ‘I don’t have a car now’)
   b. I wish I had had a car when I was a student. (conveys ‘I didn’t have a car then’ – nothing about whether I have one now)

The second environment is conditionals:

(ii) a. If he were smart he would be rich. (the condition does not hold at Present – he is not smart and he is not rich)
   b. If he had been smart, he would have been rich. (the condition did not hold at a particular time in the Past – but nothing is said about Present)
this corresponds to Progressive) and the “cul” (Culminate) notion. The former does not entail the latter: Djon sata trabesaba rua ‘Djon was crossing the street’ does not entail Djon trabesa rua ‘Djon crossed the street’. That is: just because an event holds, it does not follow that it culminates. Sometimes there is an entailment in pairs like this, due to different reasons: Djon sata pintxaba karinhu ‘Djon was pushing the cart’ entails Djon pintxa karinhu ‘Djon pushed the cart’ (in atelic event descriptions, culmination and holding are not distinct).

Thus, while to Dowty 1979 Progressive is modal, Parsons’ Progressive is non-modal and is based on events. The proposal in Landman 1992 also involves a modal approach (as did Dowty 1979), taking events in consideration (as did Parsons 1990). One crucial notion in this analysis is the one of continuation branches, which represent the different possible worlds that can follow from an event e. For instance (32a) can be followed by (32b) or (32c), etc.

(32) a. ‘John was crossing the street.’
   b. ‘John crosses the street.’
   c. ‘John was hit by a bus.’ (conveys: he did not cross the street)

Portner 1998 reformulates Dowty’s modal approach, proposing a solution to its initial problems that uses a sophisticated theory of modality, akin to the one in Kratzer 1981, 1991. The formalizations of these theories are not among the goals of the present work.

Many more examples of approaches on these crossroads could be given. Given the need of narrowing the scope of this summary according to a criterion of utility, I will only make reference to two more studies. Recall that we have hitherto been looking at the overlapping lines in Aspect: Viewpoint and Aktionsart.

As an example of the interactions between Tense, Aspect and Modality, Bhatt 1999 shows that the English modal construction was able to is ambiguous between the implicative reading ‘managed to’ and the reading ‘had the ability to’:

(33) a. John was able to eat five apples in an hour.
   b. Yesterday, John was able to eat five apples in an hour. (past episodic)
   c. In those days, John was able to eat five apples in an hour. (past generic)
The author proposes that both are derived from the underlying predicate ABLE, with the semantics of the implicative ‘managed to’, being the ‘had the ability to’ reading derived by the combination with a generic operator. In languages where genericity (Imperfective) is marked morphologically (such as French), these two readings are expressed by distinct forms. In (34) we have the worth mentioning examples in French (from Hacquard 2005):

(34)  
a. Jean pouvait soulever cette table, mais il ne l’a pas soulevée.  
‘Jean could lift this table, but he didn’t lift it.’

b. Jean a pu soulever cette table, # mais il ne l’a pas soulevé.  
‘Jean could lift this table, # but he didn’t lift it.’

From these examples (and similar pairs in languages like Modern Greek or Hindi), we can see that ability modals + Perfective Aspect entail actuality, while ability modals + Imperfective Aspect have no actuality implication. In (34b) we see that a continuation which negates an actualization to a ‘be able’ + Perfective Aspect construction results in an infelicitous sentence. In his work on this topic, Bhatt considers the same pattern for many distinct languages, and includes examples in Hindi (his own judgments) and Modern Greek (provided by Sabine Iatridou). In Hindi, the imperfective morphology in question is the one marking genericity, taa, not the one marking progressives, the auxiliary rah (as we have seen, this distinction holds in some languages, such as Hindi and in Capeverdean, but not in others, such as Modern Greek).

Thus, the difference between English and some other languages is that in the former the Perfective / Imperfective distinction for the ‘was able to’ construction is not provided by morphology on the verb, but rather by other expressions in the sentence, such as ‘yesterday’ / ‘in those days’. In Capeverdean the same distinction is marked morphologically. Similarly to Hindi, there is a distinction Imperfective Generic-Habitual / Imperfective Progressive, the one at issue here being the (Past) Habitual marker. The Capeverdean effects are similar to the ones in (34), for French. Furthermore, Capeverdean may also show an Imperfective morpheme on the second verb, giving way to a third hypothesis, not included in (34) (since the Imperfective morphology is not available in French for the embedded verb). The effects for this third Capeverdean possibility are similar to the ones in the first example.

Consider the trio in (35):
(35)  a. Djon konsigiba bebe garafa grogu d’umbes, mas e ka bebe-l / y e bebe-l. 
    ‘Djon could drink a bottle of grogu, but he didn’t drink it / and he drank it.’ 

    b. Djon konsigi bebe garafa grogu d’umbes, # mas e ka bebe-l. 
    ‘Djon could drink a bottle of grogu, # but he didn’t drink it.’ 

    c. Djon konsigiba bebeba garafa grogu d’umbes, mas e ka bebe-l / y e bebe-l. 
    ‘Djon could drink a bottle of grogu, but he didn’t drink it / and he drank it.’ 

Over the next sections we shall see that in certain constructions an embedded -ba entails counterfactuality. If this would apply here, we should expect that the actualization in (35c) would be an infelicitous continuation. This seems not the case, though. As a matter of fact, this mark of Imperfective on the verb embedded by ‘be able’-ba does not entail the negation of the actuality. This type of environments must be subject to further investigation.

In a nutshell, here is what we know so far about the above Imperfectives:

i) as expected (considering cross-linguistic evidence), Imperfective (Habitual) morphology on the modal verb (either a modal auxiliary, like pode/konsigiba ‘can’/‘be able to’, or a verb expressing a wish such as kre ‘want’, which may also contribute a modal interpretation), as in (35a), is responsible for the non-implicative reading;

ii) the addition of a given morpheme, -ba, to the lower (lexical) verb (35c) (marking which would render any actualization infelicitous for kre constructions), does not change the result for ‘be able’ constructions, which supports the specific nature of the latter.

This is shown in (36), where a) and b) are from Bhatt 1999, and c) is on Capeverdean.

(36)  a. Past (Pfv(CAN)[VP]) → managed-to (implicative – entails actualization) 

    b. Past (Impfv(CAN)[VP]) → had-ability-to (non-implicative) 

    c. Past(Impfv(CAN)[Past(Impfv(VP))]) → had-ability-to (non-implicative) 

A detailed description of more relevant data is left for the subsection 2.3.2, where I shall organize the different TMAs and discuss their values and possible combinations.
The summary in the present point has had the purpose of showing the diversity of possible proposals with respect to the complex interactions between Aspect (both in the grammatical, “point of view”, and the lexical, Aktionsart, sense) and Tense.

One common observation emerges from them all: the interdependencies that, in my opinion, point to a compositional building, involving predicates, complements, functional morphemes, adverbials of various categories and prepositional adjuncts.

In the next point, I shall summarize some recent theoretical approaches defending a different perspective on the possible relation between meaning (with respect to aspectual reference) and structure (and also the problems surrounding structural Case), including the questions raised by different approaches and proposals.

2.3.1.2 Meaning and structure

One proposal that establishes a connection between some qualities of predicates and Case is the one assuming that Partitive objects occur with unbounded (cumulative, divisible) predicates (Kiparsky 1998, a.o.). We can see this in the following English contrast, where (37b) and (37d) are ill-formed because bounded (limited) predicates do not combine with expressions like ‘a bit’ and ‘a lot’:

(37)  a. ‘to touch the picture a bit’ (unbounded)
       b. * ‘to buy the picture a bit’ (bounded)
       c. ‘to look for the book a lot’ (unbounded)
       d. * ‘to find the book a lot’ (bounded)

Although it should not be confused with telicity, the [+ bounded] feature is not available in the Present. In some languages it is associated with Perfective, whereas [- bounded] feature is associated with Imperfective. According to De Hoop 1992, Partitive is the weak Case (produces non-specific readings of the object), and Accusative is the strong Case (produces specific readings of the object). Both are checked by the verb, the Partitive in situ and the Accusative in Spec,vP. According to other hypotheses on the syntactic mechanisms involved in Case, Partitive is checked by a preposition, while Accusative is checked by the verb.
As for the mapping of complex events onto syntax, some approaches (Ramchand 2003, Rothstein 2004) propose the semantic decomposition of these events, with each of its subparts being associated with a syntactic projection. The detailed description of these proposals is beyond the objectives of the present work.  

Still in the field of trying to map Aktionsart onto syntax, the proposals in Kratzer 2004 and Borer 2005 include a syntactic projection for telicity, on which Accusative Case is dependent. Among the various reasons for Borer to refuse Krifka’s compositional proposal is that it relies on theta-role assignment, making some wrong predictions. Such as: we could expect that ‘push the cart’ and ‘push the button’ have the same value with respect to telicity, given that ‘push’ is not affected by the direct object.

This does not hold, though. Consider the equivalent Capeverdean examples:

(38) a. *Pedru pintxa karinhu na supernercado.* (atelic)
    ‘Pedru pushed the cart in the supermarket.’

b. *Djon kai pamodi Pedru pintxa-l.* (telic)
    ‘Djon fell because Pedru pushed him.’

Hence, what is involved in the atelic property of ‘push the cart’ and the telic property of the ‘push the button’ (or ‘push someone’ for that matter) is not the specific lexical contribution of ‘push’, ‘cart’ and/or ‘button’, but the knowledge that we have about life, more specifically about events like ‘push the button’.

The facts in different languages, however, do not obey the prediction that any predicate may combine with a proposed syntactic projection for telic semantics (AspQ).

This short summary has, here, the intention of showing proposals somewhat different from the one I will adopt for Capeverdean. Given the compositional strategies that in the language are clearly at stake when we try to describe the data and the temporal reference derivation, I will avoid for now any type of one-to-one mapping (one morpheme/word/expression to one distinct functional projection).

---

33 Borer’s system involves functional heads that have open values; the value of each functional head is given by an operator; the notion of operator includes f-morphs (functional morphemes), abstract head features (supported by a lexical head) and also, by an indirect relation, adverbials or discourse operators.
The alternative view that I assume in the present work is – as said before – the using of label T for all the functional pieces involved. Besides what happens in the functional domain, it is obviously necessary to keep in mind the temporal meanings involved in the rest of the sentence and also pragmatic and discourse information.

Besides having shown, in section 2.2, that adverbs are not a reliable diagnostics for a rich functional domain, I have summarized in the point 2.3.1.1, the theoretical arguments, provided in the literature, that favor a compositional approach to the temporal reference (also in the sense Aspect, both Aktionsart and Viewpoint). In the present point (2.3.1.2), I have presented a short sample of ideas that have been guiding different proposals on the possible mapping of temporal reference construction onto a range of syntactic functional projections.

In the next subsection, more Capeverdean data illustrate these complex relations; in other words, this is the empirical basis that motivates the current proposal.

2.3.2 Capeverdean TMAs: what tense is it?

Tense and Aspect are, in Capeverdean, a complex matter of cooperation between a relatively small set of morphemes, auxiliary forms and a huge amount of possible combinations with verbal semantic properties, adverbial expressions, temporal clauses and discourse information. In other words, this construction of temporal reference is typically compositional. The data and the relevant descriptive topics included in this section have been chosen in order to introduce some of these complexities.

Capeverdean data are organized as follows. In 2.3.2.1 we have sentences with various occurrences of three functional particles: the morphemes ta, sata (both preverbal) and -ba (postverbal), which are subject to adjacency requirements with respect to the verb; in 2.3.2.2 we have some details on the auxiliary sta, which can be separated from the main verb by adverbs like sempri/tudora ‘always’, participating in biclausal constructions; in 2.3.2.3, there are some remarks on the opposition statives/non-statives; and finally, in 2.3.2.4, the particle dja, which can have an adverbial function but also be an aspectual marker, displaying the particularity of being

34 I shall describe sta ‘be’ in this section, but there are also modal auxiliaries, like konsigi / pode ‘can’/’may’.
pre-subject whenever the subject is a clitic form and post-subject whenever the subject is a free pronominal or a DP.

Before proceeding, however, a methodological note is necessary. In certain cases I include empirical phenomena that are not productive in the geographical area in which I have been concentrating, the inland of Santiago Island, in the Sotavento group. Two examples are the adverbial *sempri* (mentioned in Baptista 2002, but which my informants consider a borrowing from Portuguese – their own form for ‘always’ is *tudora*), and the auxiliary *sta* (+ *ta* for the embedded verb) as a progressive marker (my informants use instead the morpheme *sata*). This option intends to address the diversity and complexities involved, and these cases are identified locally. These different markings, however, do not actually contradict the fact that, as pointed out in Chapter One (subsection 1.4.2), the debate on whether there are true varieties in Capeverdean is still far from concluded in a reasonable manner.

Given all this, we must retain two points:

i) the inclusion of alternative words/expressions/strategies that appear usually connected to other “varieties” of the language is not, obviously, exhaustive (besides these two cases, I include only a few, occasional, examples);

ii) whenever I refer to variety, it is only for labeling purposes – it does not have the intention of assuming any position on the matter.

It is a fact that most of my inland informants use *sata* to mark certain progressives and that in Praia (the capital city, also in Santiago) most people use *sta* followed by *ta*. The consequences of this distinct form on the determination of what exactly a variety is (moreover, which consequences this might have on which variety to choose when Capeverdean finally becomes an official language in the country) is well beyond the goals of the present work.
2.3.2.1 Ta, sata and the post-verbal -ba

This subsection will introduce three main TMA morphemes: *ta, sata*35, both preverbal, and -ba, a bound morpheme which affixes post-verbally. Any of them may appear as the only functional morpheme in the sentence, and they may also combine, respecting the templates in (39); this list also includes a combination of the verb with a zero morpheme. As predicted in Swift & Bonhemeyer 2002, this zero morpheme combined with stative verbs (typically atelic) cannot have a perfective reading (in Capeverdean the reading of this combination is present, hence not perfective)36; the zero morpheme combined with non-statives cannot have an imperfective reading (although we cannot say that all these are typically telic, since, as shown earlier, this is a property that varies according to the possible subgroup – activities are atelic, as states, but can be telic when a quantized object is added; achievements and accomplishments are telic) (in Capeverdean it occurs with all non-statives, triggering non-imperfective readings – the Past Perfective)37:

(39)  a) V (or V + ø )
   b) ta V
   c) V-ba
   d) ta V-ba
   e) 1) sata+V (inland of Santiago);  2) sta+ta V (other Sotavento areas)
   f) 1) sata + V-ba;  2) staba + ta V(-ba)
   g) ta staba + ta V-ba
   h) * ta sata + V
   i) * sata ta V

One previous note on this list is the following: Veiga 1995 includes al as a morpheme for marking future (uncertain future like the European Portuguese form haver + de):

---

35 Baptista 2002, on the “Sotavento varieties”, does not mention this functional morpheme, although in the inland of Santiago this particle is crucial to mark progressive forms. In other studies, this morpheme is assumed to be phonologically derived from sta ta, but check ft 40 (next page) and point 2.3.2.2 for the differences between both forms which have led me to consider them as synchronically independent.

36 The relations telicity-perfectivity and atelicity-imperfectivity are among the most basically considered when we try to articulate Viewpoint with Aktiosart.

37 I shall develop this idea – and which property is relevant for the default reading when a zero morpheme is involved – in the point 2.3.2.3).
(40)  *Dios al da-nu txuba.* (Veiga 1995:199)

_Deus há-de dar-nos chuva._ (European Portuguese)

‘[some day] God shall give us rain.’

Conjoined with a Progressive _al_ may also have an epistemic modal value, like in:

(41)  _Es ora li, nha mai al sta ta papia ku nha pai. / *al sata papia_\(^{38}\)

‘Right now, my mother must be talking to my father.’

We shall see that there are other verbal aspectual/modal markers like this. One of these is _ka_, not the negation but a morpheme/verbal unit (possibly derived from _kaba_, ‘finish’\(^{39}\)), which appears with _dja_ in certain contexts and stands for ‘just finish’ + _V_.

(42)  _Dj’-e ka faze trabadju kasa._

‘He just finished doing his homework.’

The categorial status of these units needs further investigation, though. The morphemes in (39) are the ones that may be included right away in the discussion that follows.

In some Capeverdean varieties the _sata_ form is not productive. In the Sotavento areas where _sata_ is not available the progressive form is constructed with _sta ta / staba ta_, depending on the tense (this is represented in (39) e) and f)\(^{40}\). In the São Vicente group of islands (Barlavento), the correspondent morpheme is _tita_ (or _ti ta_).

---

\(^{38}\) The fact that this _al_ combination is only possible with the _sta ta_ form, not the morpheme _sata_, is one additional argument favoring the distinct category status of both progressive forms – see note 40, below.

\(^{39}\) _Kaba_ is not _ka+ba_, the past imperfective form of _ka_, but, incidentally, a present form which might be derived from Portuguese _acabar_ ‘finish’. The past imperfective form of this aspectual verb is _kababa_.

\(^{40}\) Notwithstanding the fact that _sata_ may be historically derived from _sta ta_, there is a current distinction between them: _sata_ is clearly a functional morpheme, not a verbal form; _sta_ can be a copula or an auxiliary. In periphrastic constructions (where it is an auxiliary) _sta_ is followed by one _ta_ that marks the embedded verb (the main/lexical verb of the sentence). As a diagnostics for this distinction, we know that _sata_ cannot exhibit TMAS (say, preverbal _ta_ and/or postverbal _-ba_), like _sta_ does in (i).

(i)  _Maria staba ta/* sataba laba mo._

‘Maria was washing (her) hands.’

Moreover, _sta_ can be the only verb form (a copula), in stage-level predicate constructions; _sata_ cannot:

(ii)  _Juau sta /*sata duenti._ ‘Juau is sick.’
The reasons for the ill-formedness of combinations ta sata / sata ta (in (39) h) and i)) will become clearer in the next subsections. For the nonce let us postulate the following:

i) syntactically, there seems to be no place for these two preverbal TMA morphemes – assuming that sta, as an auxiliary, is a verbal form, not a TMA morpheme of the same type (it is included in (39) only because it will be important to establish some distinctions between the constructions in which it intervenes, with respect to temporal reference construction);

ii) semantically, it makes sense that these two preverbal morphemes cannot combine, given the contribution of each of them.

There are two other, less common, morphemes that affix postverbally: -du and -da. They mark passive constructions, present and past, respectively; hence, they encode voice plus tense. As in (43), where (43b) and (43d) are impersonal passives:

(43) a. Maria tradu di trabadju.
   ‘Maria has been fired. [she is unemployed now]’

   b. Kasas sata bendedu / * Sata bendedu kasas.
   houses TMA sell.TMA
   ‘There are houses for sale.’

   c. Maria tra da di trabadju [mas e torna podu].
   ‘Maria had been fired [but she has been readmitted].’

   d. Umbes ta badjada txeu [mas gosi ta djobedu tili vison txeu].
   ‘Once there used to be lots of dancing [but now there is lots of watching tv].’

The contrast between sata and sta ta (see footnote 40) will also be introduced in this subsection, but it will be mainly discussed in the next one, 2.3.2.2.

In order to describe the data one must use the distinction, not always clear and not sufficient most of the times, between stative and non-stative verbs. For a coherent proposal on the stative / non-stative listings of Capeverdean verbs, including a review of previous proposals in Silva 1990 and Suzuki 1994, see Baptista (2002:95-101).
For the present purposes the only information needed is that, as said earlier, to mark a
certain temporal interpretation some Capeverdean verbs require a given functional
morpheme, while others do not. This distinction is related to semantic differences
between the predicates and, for practical reasons of the present chapter, the verbs of the
first group are here called “non-statives” (with which a zero morpheme results in
perfective readings) and the ones that belong to the latter are called “statives” (with
which a zero morpheme cannot have a perfective reading – hence, the result is a present
reading, the past being marked by adding -ba, as we shall discuss below).

A few verbs, like *gosta* ‘like’, have an ambiguous behavior, since for the
combination with a zero morpheme both Past Perfective and Present readings are
allowed, depending on the other temporal information in the sentence.

My first collection of data appeared to show some other differences between *gosta*
‘like’ and *sabe* ‘know’, both considered “statives”: the former seemed to allow *ta*, the
latter did not. In more recent eliciting meetings with my informants, some occurrences
of *ta* with *sabe* have been disclosed. They do not mean the same, though. With *sabe*
‘know’, in declarative sentences *ta* always marks Future, not Present. Hence, in Future
the distinction between statives and non-statives is deactivated, for Future is also
marked by *ta* with non-statives (Future readings for non-statives are provided by the
context). This will become clearer in the next tables and the descriptions that follow.

In previous tables on this subject, I have included both “non-stative” and “stative”
verbs, in order to provide the contrasts between these groups regarding TMA selection.

Things are much more complex, though, making it rather difficult to include
every piece of relevant basic information in one table alone. Therefore, I have made the
option of separating “non-stative” from “stative” verbs: tables (44) and (45),
respectively. This does not imply considering both these groups as homogeneous – we
shall see that the group of the so-called statives is not.

Nevertheless, we must not forget the following fact, as noticed in the summary of
different theoretical observations in the previous subsection: there are common points
not yet captured by any alternative taxonomy between i) states (predicates that, without
considering any independent information available in the context, denote qualities) and
ii) events that are not states (mainly activities and accomplishments).
Both tables present simple clauses; paraphrases must be taken as comparative examples; they may be different in other specific discourse and linguistic contexts. For each reading I propose a label and add the corresponding template of the list in (39).

(44) Non-stative verbs

<table>
<thead>
<tr>
<th>Verbal aspect</th>
<th>Example (from (39))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Past Perfective (PastPer)</td>
<td><em>E odja</em> – (39) a) ‘he saw’</td>
</tr>
<tr>
<td>Present (Pres); Future (Fut)</td>
<td><em>E ta odja</em> – (39) b) ‘he sees’; ‘he will see’</td>
</tr>
<tr>
<td>Anterior Past (AntPast)</td>
<td><em>E odjaba</em> – (39) c) ‘he had seen’</td>
</tr>
<tr>
<td>Past Habitual (PastHab)</td>
<td><em>E ta odjaba</em> – (39) d) ‘he used to see’</td>
</tr>
<tr>
<td>Present Progressive (PresPro)</td>
<td><em>E sata</em> / <em>sta ta odja</em> – (39) e) ‘he is seeing’</td>
</tr>
<tr>
<td>Past Progressive (PastPro)</td>
<td><em>E sata odjaba / E staba ta odjaba</em> (39) f) ‘he was seeing’</td>
</tr>
<tr>
<td>Past Progressive Habitual (PastProHab)</td>
<td><em>E ta staba tudora ta odjaba</em> (39) g) ‘he was always seeing (when….’</td>
</tr>
</tbody>
</table>

(45) Two stative verbs

<table>
<thead>
<tr>
<th>Verbal aspect</th>
<th>Example (from (39))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present (Pres)</td>
<td><em>E gosta</em> (39)a) ‘he likes’</td>
</tr>
<tr>
<td>Future (Fut)</td>
<td><em>E ta gosta</em> (39)b) ‘he will like’</td>
</tr>
<tr>
<td>Past Habitual (PastHab)</td>
<td><em>E gostaba</em> (39)c) ‘he liked’</td>
</tr>
<tr>
<td>Conditional (Cond)</td>
<td><em>E ta gostaba</em> (39)d) ‘he would like’</td>
</tr>
<tr>
<td>Pres. Progressive (PresPro)</td>
<td><em>E sata</em> / <em>sta ta gosta</em> (39)e) ‘he is liking’</td>
</tr>
<tr>
<td>Past Progressive (PastPro)</td>
<td><em>E sata gostaba / staba ta gosta</em> (39)f) ‘he was liking’</td>
</tr>
</tbody>
</table>

Two verbs denoting states, *e* ‘be’ and *tene* ‘have momentarily’, may have in certain, very specific cases, perfective forms: *foi* and *tevi*, respectively. These correspond to Portuguese forms *foi* and *teve*.

41 With non-stative verbs, the “present” form, marked with *ta*, may have some distinct values, as expected. However, with stative verbs, only the future reading is available for the form marked with *ta*.

42 Recall from (39) that the *sata* form is only available in the inland of Santiago; *sta ta* is mainly used in some other Sotavento areas. In Barlavento areas, the progressive marker is *tita*.

43 By conditional – Cond – I do not mean a Mood, but a given reading in a sentence. Conditional readings presuppose there is a condition (antecedent) *p* and a consequent *q*, the relation between both being established by an assertion like ‘if *p* *q*’.

I have decided to name Conditional (meaning Past Conditional) this combination of morphemes *ta* and *-ba* on statives since, with these predicates, it only appears in this environment. The same combination on non-statives, although it appears in this type of environment as well, is common in PastHab readings (notice that this reading is provided by *-ba* alone in the case of statives).
Before presenting the data that will illustrate the complexities of these combinations, let us align a few somewhat superficial ideas that we can infer from (44) and (45):

A. **With non-stative verbs like *odja* ‘see’:**
   a) the bare form of the verb (absence of any TMA morpheme – or a zero morpheme) means Past Perfective (PastPer);
   b) *ta* surfaces preverbally in simple Present (Pres) constructions;
   c) *-ba*, assumed in Costa & Pratas 2004 to be affixed postverbally by lowering\(^{44}\), may mark a sort of pluperfect (AntPast), or occur in the antecedent of past conditionals (subjunctive?). It may combine with others for Past constructions: with preverbal *ta* for PastHab and with *sata* for PastPro. In a sense, thus, *-ba* marks anterior: with *ta* this anteriority is relative to Pres, without *ta* it marks anteriority relatively to PastPer.
   d) *sata* (which is exclusive of the inland of Santiago and is here assumed, against Veiga 1995, as being one word and not *sa ta*, for *sa* does not exist independently) and *sta ta* surface preverbally and mark Progressive (either PresPro, or PastPro – with *-ba*);
   e) one particular past progressive reading refers to an habitual/repeated situation in the past, which is marked in both the inland of Santiago and other Sotavento regions, with *ta staba* + *ta V*-*ba* (PastProHab).

B. **With stative verbs such as *gosta* ‘like’ and *sabe* ‘know’:**
   a) the bare form of the verb (or a zero morpheme) means simple Present (Pres);
   b) when *ta* occurs, the reading must be conditional (some condition must have been introduced), either as Future (Fut) or, if combined with *-ba* (in which case it does not build the PastHab, as is possible with non-statives), the past Conditional (Cond);\(^ {45}\)
   c) the postverbal *-ba* marks Past; the reading is imperfective (PastHab); there is no specific morphology for Past Perfective (PastPer) for these verbs;\(^ {46}\)
   d) there are differences between *gosta* and *sabe* concerning progressive forms: these are possible with former, and not possible with the latter; this one, thus, behaves as expected regarding the impossible combination with progressive morphology.

---

\(^{44}\) Lowering is the morphological operation in which a zero-level element trades its ‘head-to-complement’ relation with its complement for a relation of affixation to the structural head of its complement. See 2.4.2.1, ahead, for arguments in favor of the affixation of T to V – instead of V raising.

\(^{45}\) Past Conditional is, in fact, the Future of the Past. Hence, *ta* marks with these verbs a form of Future.

\(^{46}\) However, there are differences among them: we shall see that, in certain environments, *gosta* ‘like’ allows for a PastPer reading, while *sabe* ‘know’ never seems to allow for it
Over the next pages we are going to consider some sentences with different TMA combinations\textsuperscript{47}. A sample is provided in order to illustrate the diversity of possible constructions, each morpheme being able to contribute distinct meanings to each of them. This bunch of examples is guided by the more specific goal of supporting the main proposal in this chapter: Capeverdean TMA markers surface on adjoined functional heads with the label T. This idea will be developed in section 2.4. For each sentence a descriptive note is added.

(46) a. *Djon ka* \textit{ta} \textit{ben} \textit{ku} nos.  
*Djon NEG TMA come with 1PL*  
‘Djon does not come /will not come with us.’  

b. *Djon ka* \textit{ben} \textit{ku} nos.  
*Djon NEG come with 1PL*  
‘Djon didn’t come with us.’

The sentence in (46a) illustrates a context where \textit{ta} seems to mark Pres (Tense). Without it (the bare form of the verb, or the combination with a zero morpheme) the sentence would be Past (46b). It is not this simple, though, at least when we consider other types of predicates, such as in (47).

(47) a. *Djon ka* \textit{ta} \textit{prende} \textit{faxi}.  
*Djon NEG TMA learn quickly*  
‘Djon doesn’t learn quickly.’  

b. ??? *Djon ka* \textit{prende} \textit{faxi}.  
‘Djon didn’t learn quickly.’

The verb \textit{ben} ‘come’, in (46), refers to an achievement, naturally telic. For \textit{prende} ‘learn’, in (47), it depends. In other words, there are aspectual implications in this Tense shift, since Pres is typically imperfective and the Past form in (47b) would be Perfective. At a first glance one would need to know which thing Djon didn’t learn

\textsuperscript{47} I must recall here that the notation TMA is merely of convenience, since the main reference involved is temporal – that is, basically, of Tense and Aspect –, considering the view under which Modality is of a different nature (about possibility and necessity) and also that Capeverdean does not display any distinctive mood morphology, in the sense available, for instance, in certain Romance languages, such as French or European Portuguese.
quickly – and that is why the PastPer without a verb complement is odd. There is the immediate reactive question: *E kusé?* ‘What [is the thing that he did not learn quickly]?’

In other words, PastPer requires a lexical object for *prene* (a DP, like in *prene portuges* ‘learn Portuguese’, or a VP, like in *prene badja funana* ‘learn (how to) dance funana’), whereas Pres allows for the meaning of something like ‘Djon takes time to learn (anything)’, ‘Djon is not a good learner’.

It can likewise be used in contexts where, for instance, Djon has been cheated more than once and a friend of his comments to someone else: *Djon ka ta prende (faxi)!* Thus, without a specified complement *prene* is an activity; furthermore, *(ka) prende faxi* refers, actually, to a quality.

In other words, these effects brought by the shift from PastPer to Pres depend on the introduced imperfectivity involved in the different tenses and not on the different Tense itself; this is evident from the fact that the same effects would apply if it were PastHab – *Djon ka ta prendeba faxi*.

The line of reasoning is, then, on the close dependencies of Tense and Aspect, and on the difficulty in stating which of them *ta* specifically marks.

The implications in question will become clearer if we follow the distinction of non-stative events (as opposed to “states”) – that is, activities, accomplishments and achievements – between telic and atelic. A change of state is associated with a telic event, while it is not associated with an atelic one. This telic / atelic property is not fixed for each predicate; recall that the addition of the Direct Object ‘a mile’ gives rise to telicity for the verb ‘run’: ‘running’ (activity) is atelic and ‘running a mile’ (accomplishment) is telic (Krifka 1998). Recall also that this addition of a DO giving rise to telicity is one more argument favoring the compositionality of Aspect, not only Aspect in the grammatical sense of Viewpoint, expressed by functional morphology among other sentence elements, as said before, but also Aspect in the sense of Aktionsart, the temporal constituency of events. Making use of these concepts, and questioning which Viewpoint-Aktionsart interactions are possible in Capeverdean (some languages are more restrictive than others with respect to these interactions), we can say that the atelic construction in (47) sounds odd with a perfective like PastPer⁴⁸.

⁴⁸ Unless, of course, there is a cognate or a null object involved, like in the dialog pair on an episodic event: A. *Onti Maria ka prende badja funana* ‘Yesterday, Maria did not learn how to dance funana.’ B: *Mas Djon, e prende faxi.* ‘But Djon, he learnt quickly.’
For the moment one more note is relevant on the sentences in (46) and (47), addressing something we will return to in the present chapter: the position of the negative marker \textit{ka}. It surfaces preverbally and also pre every TMA morpheme, the only exception to this being the present copula \textit{e}, which appears in a pre-negation position (this is a puzzling phenomenon that will be analyzed in subsection 2.3.3).

In (48) two occurrences of \textit{ta} and -\textit{ba} participate in an auxiliary past construction. Both the matrix and embedded clauses illustrate what is a very productive cooperation between the two morphemes.

\begin{itemize}
  \item[(48)]
    \begin{enumerate}
      \item[a.] Mudjeris\textsuperscript{49} \textit{ta staba} (tudora/sempri) \textit{ta kantaba} (tudora/*sempri).
        \begin{tabular}{llll}
          women & TMA & AUX-TMA & always \\
          & & TMA & sing-TMA
        \end{tabular}
        ‘The women were / used to be (always) singing.’
      
      \item[b.] Mudjeris \textit{ta staba} (tudora/sempri) \textit{ta kanta*(-}\textit{ba}) (tudora/*sempri).
      
      \item[c.] Mudjeris *(\textit{ta}) staba (tudora/sempri) \textit{ta kantaba} (tudora/*sempri).
    \end{enumerate}
\end{itemize}

Although at first glance, without the adverb ‘always’ (neither in the medial nor in the final position), the sentence in (48a) could seem to have a simple progressive reading (in the sense that they were singing when I arrived, for instance), the fact is that this construction, with \textit{ta} before \textit{staba}, is only possible – that is, it is obligatory, as we can see from the notation in (48c) – in this very sense of some kind of an habitual (iterative) past progressive, obviously emphasized by the adverbial \textit{tudora/sempri}\textsuperscript{50}.

This means that at different moments, repeated in a certain period of time, the women used to be singing. In Capeverdean, this context might be given by a temporal information encoded in a PastHab adverbial clause: Like in (49):

\begin{itemize}
  \item[(49)] \textit{Kantu bu ta txigaba, mudjeris ta staba tudora ta kantaba}.
    \begin{tabular}{llllll}
      when you & TMA & arrive. TMA, women & TMA & AUX-TMA & always TMA & sing-TMA
    \end{tabular}
    ‘When you arrived (PastHab), women used to be always singing.’
\end{itemize}

\textsuperscript{49} In Capeverdean, the plural morpheme is -\textit{s}, plural being marked as a singleton, as proposed in Castro & Pratas 2006.

\textsuperscript{50} Most of my informants refuse \textit{sempri}, which seems to be a borrowing from Portuguese, and use \textit{tudora} or \textit{tudu dia} instead.
I have proposed, in the table in (44), to label this particular context as PastProHab, as a
distinction from the PastHab construction – the combination of ta V-ba – and also from
the PastPro construction – i) stabatha V(-ba); ii) sata V-ba.

Recall, still from table (44), that only “simple” progressive contexts – both in the
present and the past – have two different variants in Sotavento: the one with sata (in the
inland of Santiago island) and the one with the auxiliary sta (other geographic areas).

Back to the examples in (48), if the adverbial were not there, and maintaining the pre-
 auxiliary ta, this past “progressive habitual” (PastProHab) reading would still hold. This
reading of the auxiliary preceded by ta – which is common in the Sotavento group –
may be even more clearly marked by the sentence-initial adverb kelbes ‘in the old days’.

We will return to this distinction among past imperfective constructions. For the
moment, let us just check the contrast with (50) – an example not common in the inland
of Santiago island (cf. with (51)) – where the combination with the pre-auxiliary ta
would not be grammatical, for some other information in the sentence clearly
determines a simple PastPro reading:

(50) Kantu (ki) bu txiga, mudjeris (*ta) stababa
when (that) you arrive, women TMA AUX-TMA TMA sing-TMA
‘When you arrived (PastPer), the women were / (*used to be) singing.’

From the contrast between (48)/(49) and (50) we could infer that ta clearly marks some
specific Aspect reading, which must be added to the information from (46) – that it
marks Tense (Pres).

We must also retain from these examples one relevant information on -ba: from
the notations in (38b) we can see that -ba on the main verb is not optional in
PastProHab (always an auxiliary construction); from (50), we can see that it is optional
in the version of PastPro that is built with the auxiliary.

Interestingly enough, and as said before, only the “simple” progressive reading of the
auxiliary construction (which in the past, PastPro, is stabata V(-ba)) may also be built
with sata. The PastProHab cannot. We see this in (51a) and (51b):
Hence, we might conclude that *sata clearly marks progressive (Aspect). We shall see in PresPro constructions that this value of *sata is confirmed in the present.

In this past construction with *sata – for which, as we know by now, only the PastPro reading is available – the -ba on the verb is obligatory; its absence, that is, a *sata V version, means Pres, and this tense is disallowed in (51c), since it is incompatible with the past localization marked by kantu ki bu txiga ‘when you arrived’.

From this we could jump, again, into the conclusion that *sata marks progressive (Aspect) while -ba contributes the past (Tense) reference. This could be an interesting conclusion if there were not this contrast between (48) and (50), where -ba is obligatory in the former (check (48b)) and optional in the latter.

Since these two contexts are distinguished by a habitual property (it holds for (48) and does not hold for (50)), we must consider that this (Aspect) feature is somehow related to the obligatoriness of -ba – this means that this morpheme cannot only mark past (Tense).

Moreover, the combination in (52) is never possible, nor in habituals or progressives, hence it is not clear why -ba – if it were exclusively a tense marker – would be allowed when there is a -ba on the auxiliary, and prohibited in (52).

\begin{enumerate}
\item \textbf{Mudjeris sta ta kantaba.}
\end{enumerate}

Whereas it is possible that -ba on the verb marks past (Tense) in past auxiliary constructions – *sata is not a verb form, hence it cannot bear TMAs –, the truth is that Tense is supported by -ba on the auxiliary, not by the one on the main (lexical) verb. This last -ba must, then, have some Aspect contribution, at least for the habitual (iterative) reading, otherwise it would be difficult to explain why it is obligatory in (48b). Let alone the impossibility in (50), which may be due to the fact that the Aspect
properties implicated by the main verb + -ba are not compatible with a present Tense (given by the absence of -ba on the auxiliary).

We shall see that for instance perceptive verbs select for a progressive form of the lower verb (this will be analyzed in detail in Chapter Six, on embedded non-finite constructions). Curiously, though, these perceptive past constructions do not require an obligatory -ba on the lower verb; just as it happens with the non-iterative past progressive with sta. This distinction reinforces the necessary contribution of this lower -ba in certain particular environments.

We may confirm this in (53).

(53) a. N ta odja Maria ta badja (kinta fera di noti).
    1SG TMA see Maria TMA dance
    ‘I see Maria dancing.’ (habitual, like with ‘every Thursday.’)
b. N ta odjabo Maria ta badja(ba) (kinta fera di noti)
    1SG TMA see.TMA Maria TMA dance.TMA
    ‘I used to see Maria dancing.’
c. * N ta odja Maria ta badjaba.

The contribution of the embedded -ba, again, is not compatible with present, as seen in (53c).

The sentence in (54) is very interesting, since -ba and ta surface in different clauses: in the matrix and in the embedded clauses, respectively.

(54) N ka sabeba ma Ferro y Gaita ta toka na Praia oxi.
    1SG NEG know.TMA COMP Ferro y Gaita TMA play PREP Praia today
    ‘I didn’t know that Ferro y Gaita is playing /going to play in Praia today.’

In the matrix clause, -ba marks the past for the verb sabe ‘know’ (a stative verb). Recall that the bare form of this verb would mean Pres (the bare form of a non-stative verb means PastPer). Crucially, the PastHab reading of non-stative verbs, like kume ‘eat’, kanta ‘sing’, etc., involves the cooperation between ta and -ba; here, with a stative verb – hence, inherently imperfective –, ta is not needed, and this favors its imperfective marker status. This holds for present and past tenses.
In the embedded clause in (54), introduced by the complementizer *ma*, *ta* seems to mark a form of Pres with a (near) future (Fut) reading assured by the adverbial *oxi* ‘today’. Without this adverbial, the meaning of the embedded clause could be an habitual (iterative), like ‘I didn’t know that Ferro y Gaita play in Praia (regularly)’. Hence, *ta + oxi* assures that something that is going to happen in a somewhat posterior time with respect to the reference time (in this case, with respect to the utterance time as well). So, what is exactly the contribution of this *ta*?

I assume here that while the inherent imperfectivity of stative verbs is semantically encoded in the verb, the inherent irrealis property of future can only be encoded in some mark of Fut. In this case, this mark of Fut (or posteriority with respect to Pres) must be *ta*, with the adverbial *oxi* contributing the proximity part. Hence, and again, does *ta* mark Tense or Aspect?

In (55) all that is relevant for now is that the matrix clause is PastPer (no *ta* preceding the non-stative verb; no affixed -*ba*) and in the embedded clause we have again the morpheme *ta* marking a posterior tense.

(55)  

a. *Pedru fla Maria m’-es *ta* badjaba noti manxe.*  
   Pedru tell Maria COMP-3PL TMA dance.TMA night dawn  
   ‘Pedru told Maria they would dance all night long.’

b. *Pedru fla Maria m’-es ta badja noti manxe.*  
   ‘Pedru told Maria they will dance all night long.’

c. *Pedru fla Maria m’-e badja noti manxi.*  
   ‘Pedru told Maria that he has danced all night long’

From (55) we may list some observations:

- in (55a) and (55b) the reading of the embedded clause points to an action (whose duration limits are given by the adverbial expression *noti manxe* ‘all night long’) posterior to the one of the matrix clause, which is PastPer;
- in (55a) -*ba* contributes a past reading with respect to the utterance time;
- in (55b) the embedded event tense is posterior with respect to both the utterance and the reference time (the matrix time);
the irrealis (both in the Cond in (55a) and in the Fut in (55b)) – which seems provided by *ta*, while -*ba* marks past (the combination of both occurs in conditional environments51) – is also reinforced by the cooperation between TMA morphemes and the semantic implications of the verb *fla* ‘say’ (we do not know if something that someone said would do (55a) or is going to do (55b) has actually been done or is going to be done).

Confront this with the result in (55c), where the absence of *ta* in the embedded clause clearly induces a PastPer reading, with the correspondent actuality implications.

In fact, even the meaning of the verb *fla* suffers some changes with this, from something more similar to ‘promise’, in (55a) and (55b), to the simple ‘tell’ (in the sense of ‘inform’) in (55c)52.

Notice that if the embedded 3pl pronominal refers to Pedru and Maria, for the ‘promise’ or ‘invite’ reading to obtain the embedded clause must be introduced by the preposition *pa*, in which case the embedded *ta* would be barred53 – *Pedru fla Maria p’-es (*ta) badja(ba) noti manxe.*

Nevertheless, nothing prevents the *ma* clauses, such as we have in (55a) and (55b), from referring to some people not including Maria, in which case the reading of *fla* as ‘promise’ (not as ‘invite’, though) is perfectly acceptable.

The lack of -*ba* in the embedded clause which marks the difference between (55a) and (55b) results in the shift from Cond to Fut54; from (55b) to (55c) the removal of *ta* in the embedded clause results in the shift from Fut to PastPer, and the respective actuality implications are connected with a different meaning of the matrix verb *fla*.

---

51 Recall that Cond stands for a type of reading, not the Mood (whose existence is subject to debate).

52 The different meanings of *fla*, which have consequences on the complementizer selected as the introducer of the embedded clause, will be at issue in Chapter Six.

53 Probably because the embedded sentence is rendered infinitive.

54 We shall see that, in Capeverdean, only in counterfactual environments is the form correspondent to ‘would’ – *ta* + V-*ba* – used to refer to a time posterior both to the reference time and the utterance time (like in ‘Peter said he would take the plane tomorrow’); that is, in this case we could be able to add an ending such as ‘but I believe he won’t’. Otherwise, in non-counterfactual environments, the embedded future is used, such as in (55b).
A more evident conditional environment (one typical ‘if \( p, q \)’ relation) is built in (56), whose matrix clause corresponds to the European Portuguese past subjunctive se eu ficasse:

\[(56) \quad \text{Si } N \ fikaba \ \text{un anu na Kauberdi,} \quad N \ \text{ta papiaba kriolu dretu.} \]

if 1SG stay.TMA one year in Cape Verde, 1SG TMA speak.TMA kriolu right

‘If I stayed for a year in Cape Verde, I would speak creole perfectly.’

Do we still have any evidence to consider that \( ta \) marks only aspectual reference and -\( ba \) only tense? Or should it be the other way round?

In (56), both -\( ba \) alone (in the matrix clause) and the combination of \( ta \) and -\( ba \) (in the embedded clause) seem to contribute imperfectivity (recall that if \( ta \) were sufficient to bring this imperfective reading we would not really need -\( ba \), since the past role should be assured by the morpheme zero, as it is in PastPer with non-statives).

Moreover, if we associate \( ta \) with generic/habitual and/or some posterior notion (including some Pres constructions), while -\( ba \) marks past, what to say about this sort of past subjunctive, in the matrix clause, which is also marked by -\( ba \)?

And how come that the combination of these morphemes (typically PastHab) also means, in certain contexts, some sort of conditional, bringing along all the inherent aspectual properties?

Even though it is true that in some other languages past imperfective (indicative mood) appears in subjunctive and/or conditional contexts, the crucial question remains:

- which of these morphemes stands for Tense, Mood and Aspect?

Let us take a look at counterfactual environments in more detail, in the line of the analysis for Modern Greek proposed in Iatridou 2000. Consider, first, the English examples in (57):

\[(57) \quad \text{a. If he had taken the syrup, he had gotten better. (past Counterfactual)}
\]

\[\text{b. If he took the syrup, he would get better. (Future Less Vivid, FLV)}\]

\[\text{c. If he takes the syrup, he will get better. (Future Neutral Vivid, FNV)}\]

77
In (57a) the speaker believes that ‘he’ has not taken the syrup in some time in the past, and it is perceived that, because of this, ‘he’ didn’t get better.

The counterfactuality in (57b) is not an entailment, but an implicature – to the speaker, the actual world is more likely to be ~p; in other words, although there is still a possibility that ‘he takes the syrup’ in the future, it is more likely that ‘he’ doesn’t. It refers, thus, to the future, in spite of its past morphology (Iatridou calls this fake past, or fake tense). This is not unexpected, though, for in modal environments (where subjectivity plays a crucial role), past morphology may indicate the sense of remoteness, not a past tense event.

The sentence in (57c) clearly refers to the future, and might be given as an instruction to a caretaker; (57b) might be so, as well, with the difference being in the speaker: in the latter, for him the most likely is ~p; in the former, there is no such belief.

Returning to the Capeverdean configuration in (56), we may say it corresponds to a Future Less Vivid (FLV) conditional. Consider the Modern Greek examples in (58), from Iatridou 2000:

(58) a. An pari afto to siropi θα γίνει kala.
   if take.NONPAST.PER this syrup FUT become.NONPAST.PER well
   ‘If he takes this syrup he will get better.’

b. An eperne afto to siropi θα γίνοταν kala.
   if take.PAST.IMP this syrup FUT become.PAST.IMP well
   ‘If he took this syrup he would get better.’

If one considers, at least in these particular contexts, -ba as Past Imperfective and ta as Future, the morphology involved is similar to what we find in Modern Greek in (58b):

Iatridou contends that the imperfective in (58b) is a fake tense and a fake aspect – a fake past, and also a fake imperfective –, since the event is interpreted perfectively and might occur in the future.

Its perfectivity is confirmed by the possible combination with completive adverbials (the prepositional adjuncts of the type ‘in one minute’, as mentioned earlier – recall that these contrast with the type ‘for one minute’; the latter are the ones truly compatible with imperfectives).
Consider the Capeverdean sentence:

(59) \textit{Si e koreba faxi e ta txigaba antis di nos fepu/nos dianti.}
\hspace{1cm} if 3SG run.TMA quickly 3SG TMA arrive.TMA before of us all / POSS.1PL in.front
\hspace{1cm} ‘If he run quickly he would arrive before us all.’

The arriving can only be interpreted as perfective (it is certainly not habitual), and nothing determines that its time location cannot be in the Future. Thus, the Past Imperfective of the consequent is \textit{fake} Aspect and \textit{fake} Tense.

Now consider one European Portuguese Future Less Vivid (FLV) context, where Past Imperfective (indicative mood) is productive in the consequent (60a) as well, in spite of a conditional form being available (60b). Both sentences may have a Future reading.

(60) \hspace{0.5cm} a. \textit{Se ele corresse depressa, chegava antes de nós todos.} \\
\hspace{1cm} if 3SG run.SUBJ quickly arrive.PAST.IMP before of us all \\
\hspace{0.5cm} b. \textit{Se ele corresse depressa, chegaria antes de nós todos.} \\
\hspace{1cm} if 3SG run.SUBJ quickly arrive.COND before of us all

In this case, curiously, we might say that in (60a) we have a \textit{fake tense}, a \textit{fake aspect} and also a \textit{fake mood}.

The crucial difference from Capeverdean is the following: whereas in European Portuguese the Future involved in conditionals is only encoded morphologically in (60b), in Capeverdean it is encoded in a regular Past Imperfective morphology for non-stative verbs, that is, the correspondent to the Portuguese form in (60a).

Resuming the counterfactual environments overview, departing from Iatridou 2000, the \textit{fake imperfective} and the \textit{fake past} involved in FLV counterfactuals may occur in a different type of environment as well: the \textit{wish} counterfactuals. In Modern Greek, these are constructed with the verb ‘want’, and a particular set of morphological markings: the matrix clause has the same morphology as the consequent of a FLV counterfactual, namely for future, imperfective and past (although in the notations the morpheme that is identified as future in FLV counterfactual consequents, is identified as modality in the matrix of the \textit{wish} counterfactuals).

79
This example (from Iatridou 2000:239) conveys that Kostas does not drive a red car. As for the morphology in the embedded clause, the distinction between past and non-past is crucial: past participates in counterfactual wishes (the ones conveying that the actual world is not how the subject would like it to be); non-past participates in non-counterfactual wishes (these express a desire possible to fulfil – in this case, it is more like ‘want’ than ‘wish’). The contrast is illustrated in (62):

(62) a. θα ιθελα να ιταν πσιλος.
    MOD want.PAST.IMP NA be PAST tall
    ‘I wish he were tall.’

b. θα ιθελα να ινε πσιλος.
    MOD want.PAST.IMP NA be NONPAST psilos
    ‘I would like him to be tall.’

This description is essential for some similar ‘wish’ contexts in Capeverdean that we will see later (the relevant relations and examples will be recaptured then).

For the time being, it is important to notice that, still in Modern Greek, it is not the morphology on ‘want’ (identical to the FLV counterfactual consequent, although, as said before, the morpheme that is identified as future in this is identified differently in

---

55 The status of this Modern Greek na has been subject to debate. According to Iatridou 1993 [1988], there are no embedded infinitival clauses in Modern Greek, and this na may be more like a modal marker, followed by a predicate marked [+ T] – this is proved by the possible marking for past. In Iatridou 2000, however, the possible ambiguousness of na is discussed – na clauses may be infinitival, subjunctive (if this exists in the language) or ambiguous between the two (Iatridou 2000:fn 37). For this author, in this paper, though, what is relevant is that there is no reason to believe the subjunctive is involved in counterfactuals morphology in Modern Greek.

In a different perspective, Bošković 2004b considers that na is a marker for subjunctive, introducing for instance negative imperatives, such as (iic).

(i)  a. Diavase!
    read.IMP

b. *Den/mi diavase!
    NEG read.IMP

c. Na mi diavazis!
    SUBJ.MARK NEG read.SUBJ

---
‘want’ clauses) that determines the counterfactuality in wish contexts, but rather the past morphology on the embedded verb.

Iatridou’s proposal on the fake past intervening in counterfactual environments goes as follows: the morpheme in question has not a temporal past meaning as a primitive; this meaning is derived from the combination of the morpheme with other interpretive elements available.56

As for the fake aspect, although in Modern Greek there is no morphological distinction between the imperfective with a progressive reading and the imperfective with an habitual/generic reading, Iatridou 2000 shows that in counterfactuals what is really at stake is fake habitual (progressive is only involved when the speaker is indeed referring to an ongoing event).

This is confirmed by the data in Hindi (where past generic/habitual and past progressive are marked with different morphemes). And, as we can observe in (63), this is so in Capeverdean as well.

(63)  
   a. *E ta kumeba karni, mas gosi e ka ta kome mas.*
       ‘He used to eat meat, but now he doesn’t eat anymore.’

   b. *Si e stabu ku fomi, e sata kumebu.*
       ‘If he was hungry he would be eating.’

The opposition between a fake tense (corresponding to English ‘would’ – a Future Less Vivid, FLV) and a real tense (corresponding to English ‘will’ – a Future Neutral Vivid, FNV) depends – in Modern Greek as well as in Capeverdean – on the presence/absence, respectively, of the past habitual morpheme, both in the antecedent and the consequent.

56 I will not go into the detailed formalization about past tense and how it establishes a relation between a topic time (or the reference time: the time we are talking about) and the utterance time, and not between the utterance time and the situation (event) time. What matters for now is the relation of exclusion – the topic time excludes the utterance time. This also leaves a crucial independency (that is, no anteriority / posteriority pre-established) between utterance time and event time – the time relation between these is given by some elements in the sentence other than the past tense morphology. Consider the following:

(i) *Si e stabu ku fomi, e sata kumebu kantu N twiga.*
   ‘If he was hungry he would be eating at the moment I arrived.’

Furthermore, the above mentioned relation of exclusion between the topic time and the utterance time is reflected on the types of conditionals according to the Aktionsart of these: a telic predicate (the result is an FLV); an individual-level stative predicate (the result is a Pres counterfactual); a stage-level stative predicate (the result is either an FLV or a Pres counterfactual). See details of this in Iatridou 2000:250.
Check the following contrast:

(64) a. Si e koreba faxi e ta txigaba nos dianti. (FLV)
   if 3SG run.TMA quickly 3SG TMA arrive.TMA POSS.1PL in.front
   ‘If he run quickly he would arrive before us.’
b. Si e kore faxi e ta txiga nos dianti. (FNV)
   if 3SG run quickly 3SG TMA arrive POSS.1PL in.front
   ‘If he runs quickly he will arrive before us.’

For Modern Greek, Iatridou uses the notation NONPAST.PER for the absence of these relevant markings. The same could be used for Capeverdean. Recall the Modern Greek examples in (58), here repeated as (65) (the order of the examples is reversed, to match the correspondent contexts in (64)):

(65) a. An eperne afto to siropi θa γι' inotan kala.
   if take.PAST.IMP this syrup FUT become.PAST.IMP well
   ‘If he took this syrup he would get better.
   b. ’An pari afto to siropi θa γι' ini kala.
   if take.NONPAST.PER this syrup FUT become.NONPAST.PER well
   ‘If he takes this syrup he will get better.’

Let me resume the contrast between the embedded clauses in (55), here repeated as (66).

(66) a. Pedru fla Maria m’-es ta badjaba noti manxe.
   Pedru tell Maria COMP-3PL TMA dance.TMA night dawn
   ‘Pedru told Maria they would dance all night long.’
b. Pedru fla Maria m’-es ta badja noti manxe.
   ‘Pedru told Maria they will dance all night long.’
c. Pedru fla Maria m’-e badja noti manxi.
   ‘Pedru told Maria that he has danced all night long’

In (66a) the embedded -ba has clearly the meaning of past, since its elimination results in a change of tense, loosing the past reading. As said earlier, if we add a counterfactual information – such as ‘but they will not be able to’ – the embedded Past Imperfective is fake. Without this counterfactual information, this Past Imperfective has a plain reading
– as a Future with respect to the reference time, and Past with respect to utterance time. This is so because the matrix clause does not introduce, per se, any condition. It may be merely a statement, with no other implicatures. This seems non-problematic.

Now consider the possibly problematic part: suppose ta was responsible for the habitual Aspect in (66); this could seem coherent with the result in (66c), where there is no ta and the reading is Past Perfective (badja is, here, an activity – and activities share some common features with stative predicates –, but its selectional properties with respect to TMAṣ are typically non-stative). In (66b), however, the reading is clearly Future, not with respect both to the reference time and the utterance time.

Hence, we might conclude, instead, that in these contexts ta marks Future and it is -ba that marks Past. Thus, in (66a), ta would mark the Future with respect to the reference time, and -ba would mark the Past with respect to the utterance time. In this case, no Habitual feature seems at stake.

In a nutshell, let me recapture some ideas and doubts that have been raised so far.

We have seen that:

A. … ta and -ba combine for PastHab (one form of past imperfective) with non-stative verbs in declarative contexts (suppose ta would mark Aspect and -ba would mark Tense).

B. … -ba is the only morpheme needed to mark PastHab with statives (suppose the relevant – imperfective – Aspect properties are encoded in these verbs and -ba marks Tense; apparently this could be expected).

C. … ta is needed to mark Pres (Tense) with non-statives (suppose we would argue that present is imperfective, and with non-statives, as inherently perfective, this present imperfectivity must be derived; in this perspective, ta would be a marker for Aspect anyway).
From the observations in A, B and C, everything seems to articulate with one well-behaved description: *ta* marks Aspect; *-ba* marks Tense.

Nevertheless… we have also seen that:

D. … **PastPer** is marked by a **zero morpheme** (the bare verb form) with **non-statives**; hence, the adding of *ta* changes the sentence **Tense** (no matter the aspectual properties of this obtained Tense)\(^{57}\).

E. … in **PastProHab**, assuming *-ba* marks Tense on the auxiliary, *-ba* on the main (lexical) verb must bring some **Aspect** contribution, since it is obligatory in these environments whereas in **PastPro**, also in the version with the auxiliary, it is not.

F. … the combination of *ta* and *-ba* mark the consequent in a conditional counterfactual (**Cond**), whereas *-ba* marks the **past** antecedent (this is marked by subjunctive mood in many languages); in these contexts, the **future/posterior** feature seems to be on *ta* and both the past and the imperfective/habitual seem to be on *-ba*.

G. …if one assumes that *ta* marks Aspect, one has to decide which **value** it is: imperfective, as for **Pres**? Or posterior, as for **Fut** and for **Cond**, among others?

H. …if only *ta* marks **Aspect** (and, thus, *-ba* marks **Tense**), we must explain why *-ba* is needed in **PastHab** with non-statives, since with these verbs a past reading is provided by a zero morpheme.

\(^{57}\text{It is not my intention to plunge into a dissertation on what Tense is. Are Present or Future really Tenses? And what about Past, does it only apply to the PastPer reference, since all the others are truly a huge complex of aspectual and modal information? What about the English Perfect, whose aspectual implications are the subject of huge discussions and studies? The unstable (volatile) nature of Tense is even more acute when we look at Capeverdean data. This could seem an argument against my main proposal on the functional domain, but I believe this is actually a strong argument for it.}\)
In the table in (67) we may see more clearly the possibilities for ta or -ba when intervening alone, the combination of the two and also the possible bare form of the verb, identified by zero ø. Notice that in the first two lines, -ba and ta are considered as when they are not combined with other morphemes in the same clause.

(67)

<table>
<thead>
<tr>
<th>Tense</th>
<th>Grammatical Aspect (Viewpoint)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Past</td>
</tr>
<tr>
<td>-ba</td>
<td>✓</td>
</tr>
<tr>
<td>ta</td>
<td>No</td>
</tr>
<tr>
<td>ta + -ba</td>
<td>✓</td>
</tr>
<tr>
<td>zero - ø</td>
<td>✓(^{62})</td>
</tr>
</tbody>
</table>

The possible unexpectedness of some combinations is the reason for so many footnotes pointing out in which precise contexts each value is available.

The previous descriptive points and the table above – leading to an ultimate conclusion on the values of each morpheme that might be presented as “it depends” – illustrate how inglorious it could be to project a specific functional head to accommodate each of these morphemes.

---

58 In the antecedent of FLV conditionals, if we consider -ba as marking a fake Tense and Fake aspect this is because the reading may be Future and Perfective, instead of Past Imperfective. Recall that it could also be argued to be a fake Mood (Subjunctive, instead of Indicative).

59 In Pres for non-statives.

60 In embedded infinitive clauses selected by auxiliary constructions with sta or by perceptive verbs.

61 In the consequent of FLV conditionals, the true Tense reading may be Future, with respect both to the reference time and the utterance time.

62 With non-statives.

63 With statives.

64 In Pres for statives.
This is so for two reasons:

i) each of these morphemes may mark different values according to its environment, and even cumulate two features that are not inherently related to one another;

ii) some values are derived, obtaining by the interaction between two of them, as well as by the Aktionsart of predicates.

In other words, any proposal of a hierarchy of functional projections, each of them encoding temporal pieces of information such as Tense, Mood or Aspect, would not be faithful neither to the data nor to any concerns on clarity and explanatory adequacy.

In the next point I will explore in greater detail the different occurrences of the morphemes involved in progressive constructions: sta and sata. Recall that only in the inland of Santiago island sata is the morpheme marking some of these constructions; in this same area, the auxiliary sta is only required in the past (marked with preverbal ta and postverbal -ba) in what I have called PastProHab environments.

### 2.3.2.2 More on sta and sata

In the previous point we have seen that the sata form, used in the inland of Santiago Island, occurs in constructions with the progressive reading (both PresPro and PastPro). A “past progressive habitual” reading, however, requires the sta version (an auxiliary construction), with the relevant TMA morphemes (which sta, as a verbal form, can bear): pre-auxiliary ta and post-auxiliary -ba, both obligatory in these contexts. The ta that follows this set of ta stabar is a marker for the main (lower) verb of the sentence, since we have here, as proposed in Baptista 2002, a biclausal construction (a TP embedded in another TP, which also allows for an adverb to occur between the two functional domains).

In (68)-(71), some more examples with sata are presented, all of them in the present (PresPro), contrasting with the same sentences marked by ta. We shall see that only the

---

65 I assume that other materials generated under the label T – like, as we shall see in 2.3.3, the negation head ka – do not appear in this intermediate position for semantic reasons (meaning incompatibilities) and not because of the lack of a syntactic position.
first three examples have truly habitual (ta) and progressive (sata) readings, with the last one, (71), being a different instance of present.

(68)  
- **a. (tudora) N ta fla-u pa-u bai undi-m.**
  
  1SG TMA tell/say-2SG PREP-2SG go where-1SG.
  
  ‘I tell you (all the time) to go visit me.’

- **b. N sata fla-u pa-u bai undi-m.**
  
  1SG TMA tell/say-2SG PREP-2SG go where-1SG.
  
  ‘I’m telling you to go visit me.’

(69)  
- **a. Manel ta djunta dinheru p’-e bai Portugal.**
  
  Manel TMA save money PREP-3SG go Portugal
  
  ‘Manel saves money (all the time) to go to Portugal.’

- **b. Manel sata djunta dinheru p’-e bai Portugal.**
  
  Manel TMA save money PREP-3SG go Portugal
  
  ‘Manel is saving money to go to Portugal.’

(70)  
- **a. N ta skrebe nha primu (*karta / kartas).**
  
  1SG TMA write POSS.1SG cousin letter / letters
  
  ‘I write my cousin (*letter / letters).’

- **b. N sata skrebe nha primu karta.**

  1SG TMA write POSS.1SG cousin letter
  
  ‘I’m writing my cousin a letter.’

(71)  
- **a. Parse-m ma Suzana ta kasa.**
  
  seems-1SG COMP Suzana TMA marry
  
  ‘It seems (to me) that Suzana is going to marry.’

- **b. Parse-m ma Suzana sata kasa.**
  
  seems-1SG COMP Suzana TMA marry
  
  ‘It seems (to me) that Suzana is getting married / going to marry.’

---

---

66. This predicate, being ditransitive, requires the order of a Double Object Construction in Capeverdean (see the last section of Chapter Three, on DOCs), which is Verb-Indirect Object-Direct Object. This holds for PastPer – *N skrebe nha primu karta* ‘I wrote my cousin a letter’ –, as well as for the progressive form in (50c’). The preposition construction (the so-called to-dative) – *N sata skrebe karta pa nha prima* ‘I am writing a letter to my cousin’ –, besides being much less productive, brings a slightly different interpretation, as we can see in detail in Chapter Three.
In (68) and (69) the relevant constructions are in the matrix clauses (the embedded clauses seem non-finite). In (70) there is no embedded clause. In (71) ta / sata appear in the embedded clause (notice that in Capeverdean the subject does not raise past the verb parse ‘seem’: *Suzana parse ta kasa).

The four contexts illustrate different values of present (habitual/generic and progressive), with two versions for each of them.

At first sight, however, we can infer among the PresPro versions some distinct temporal meanings. Let us consider them:

i) An instant coincident (or immediately anterior, if we understand this as something like ‘I just told you’) to the utterance time (68b).

ii) An interval that includes a period anterior to the utterance time and extends to a period after it (69b).

iii) In (70b) we also have a notion of a time span, although we feel that this one is much shorter.

iv) As for (71b), besides this progressive meaning (which would be possible, although not so common), the most likely reading is that we have a moment – we do not know exactly which – in the near future where something is going to take place (which holds for the example in (71a) as well).

Elaborating a little further on (71), we may not have a time span for the event we are referring to – the act of getting married –, which is actually an instantaneous / episodic event, located in a (near) future, but a notion of the time span linking the reference time to the event time.

The reference time in (71a) and (71b) may not be coincident with the utterance time, since the sentence would have exactly the same shape if we added a temporal (future) expression, like oki si kretxeu ben di Fransa ‘when her fiancé has arrived from France’.
In order to obtain a past interpretation with the progressive readings of the first three examples the verb should be marked with the -ba affix. As we have seen in the previous section, these past progressive readings may also be built, in other areas of Sotavento, with staba ta, followed optionally by the -ba affix on the main verb. Recall that sata, not being a verbal form, cannot exhibit TMAs, a possibility displayed by sta (which confirms the verbal nature of the latter).

The results of transferring into PastHab the sentences of (68a), (69a) and (70a), which had a PresHab reading, are presented, respectively, in (72a), (73a) and (74a).

The previous PresPro versions in (68b), (69b) and (70b), are, respectively, in (72b), (73b) and (74b)67:

(72) a. (tudora) N ta flaba bo pa-u bai undi-m.
always 1SG TMA tell.TMA 2SG PREP-2SG go where-1SG.
‘(all the time) I used to tell you to go visit me.’

b. (Kantu ki e txiga) N sata / staba ta flaba bo pa-u bai undi-m.
‘(when he arrived) I was telling you to go visit me’

(73) a. Manel ta djunta dinheru p’-e bai Portugal.
Manel TMA save.TMA money PREP-3SG go Portugal
‘Manel used to save money to go to Portugal.’

b. Manel staba ta / sata djunta dinheru p’-e bai Portugal.
‘Manel was saving money to go to Portugal [but he got fired and gave up].’

(74) a. N ta skrebeba karta…
1SG TMA write.TMA letter…
‘I used to write a letter…’

b. [kantu bu txiga68], N staba (*tudubes) ta / sata skrebeba karta….
‘[at the moment you arrived], I was (*always) writing a letter…’

67 The PastProHab reading of these constructions, as we also have seen, is only possible with ta staba ta, with the obligatory -ba on the main verb.

68 If instead of the expression Kantu ki... we used Oki..., both meaning ‘When…’ / ‘At the time…’, the temporal reference of the sentence would switch from the Past episodic into the Future or a possibly iterative action: Oki bu txiga... ‘When you (will) arrive, I will…’ Hence, also this temporal words, with (apparently) the same function in the sentence, denote distinct tense/aspect information.
As for the sentences in (71a) and (71b), the temporal correspondences are not so linear. Let us check some possibilities in (75).

(75)  
a. Parse-m ma Suzana ta kasaba. (Cond)  
seems-1SG COMP Suzana TMA marry.TMA  
‘It seems (to me) that Suzana would marry.’
b. Parse-m ma Suzana stabat ta kasa(ba). (PastPro)  
seems-1SG COMP Suzana AUX TMA marry.TMA  
‘It seems (to me) that Suzana was getting married [when he arrived].’
c. Parse-m ma Suzana sata kasaba. (PastPro / Cond)  
seems-1SG COMP Suzana TMA marry.TMA  
‘It seems to me that Suzana was getting married [when he arrived].’
d. Parse-m ma Suzana stabapa kasaba. (past prediction)  
seems-1SG COMP Suzana BE.TMA PREP marry.TMA  
‘It seems (to me) that Suzana was to get married.’
e. Parse-m ma Suzana ta stabata kasaba. (PastProHab)  
seems-1SG COMP Suzana TMA AUX TMA marry.TMA  
‘It seems (to me) that Suzana was getting married [all the time].’

Let us now check some ideas about (75):  
i) In (75a) we observe the expected shift from a future reading to a past conditional reading: we obtain a similar context to the consequent in FLV counterfactuals, where the combination of the Future and the Past Habitual features results in a ‘would’ construction, which might indeed happen in the Future (hence, it is a fake Past Imperfective).

ii) In (75b) we only have the progressive reading, with *kantu ki*… being the only possible temporal reference expression (and not *Oki*…, which preferably points to an eventual occurrence in the Future).

iii) In (75c), with *sata*, we have both the progressive and future readings (better defined by the type of temporal expression added). For the progressive (as also in (75b)), take the following scenario: it seems that
Suzana was in the act of getting married (she was in the church with her groom, and they were pronouncing their vows) when her previous fiancé arrived in town.

iv) The possible sentence in (52d) leaves no place for any readings other than the one in the posterior/near future with respect to the reference time.

v) The sentence in (75e) is only included here to illustrate, again, the contrast between the PastProHab, with \( ta \text{ stab}a + ta \text{ V}-ba \), and one of the possible versions of PastPro, with \( \text{stab}a + ta \text{ V}(-ba) \), as in (75b).

A different type of information needed for the interpretation of \( sata \) to be correct is the one encoded in the verb semantics – that is, Aktionsart: the aspectual properties of different events (well beyond the distinction between statives and non-statives), and their interaction with some adverbials.

Let us take a very productive context: \( sata \) with a weather verb, as in (76), and \( sata \) with a telic predicate, as in (77):

(76) \text{Oxi sata txobe.}

‘Today it is raining.’ / ‘It is raining.’

(77) a. \text{Djon sata txiga oxi.}

‘Djon is arriving today.’

b. \text{Oxi Djon sata txiga.}

‘Today Djon is arriving.’

c. \text{Djon ta txiga oxi.}

‘Djon arrives today.’

d. \text{Oxi Djon ta txiga.}

‘Today Djon arrives.’

Whereas in (76) we may have a true progressive reading (an event that lasts through a time span, with the instant of the utterance being coincident with some point within this time span), in (77) the reading is never progressive in this sense. This is expected, for
the predicate denotes an achievement, impossible to be considered an on-going event. Therefore, the meaning of the sentences with *sata* is not distinct from the ones with *ta* (77c) and (77d).

What we know about (77a) and (77b) is that Djon is doing something, probably travelling here (this might be the true process implicated in the sentence) such that, at a certain moment of today, he will arrive. This is different from *Oxi sata txobe* ‘Today it is raining’, which may indeed refer to the on-going event of raining. In some circumstances, however, the latter may also denote a strong prediction: someone looks at the heavy clouds in the sky and makes this assertion as if it were true already.

Still in (77), the distinct word orders – the sentence-intial or the sentence-final adverb – have effects on the information structure: in (77a) and (77c) the adverb is focused; in (77b) and (77d), the focused part is that Djon is arriving.

Now let us take, again, a counterfactual with *sata*. We have seen that in languages where there is a morphological distinction between Habitual and Progressive (other languages, like Modern Greek, only have one morpheme, generically marking Imperfective), it is usually the Habitual morpheme that shows up in these conditional environments.

We have also seen that the Progressive morpheme in the consequent, for instance in Hindi and in Capeverdean, is limited to contexts where the speaker is, in fact, referring to an ongoing event.

As for the antecedent, and following the above mentioned possible future readings of Capeverdean *sata*, in (78c) the template *sata V-ba* may yield the same consequences as *V-ba*\(^{69}\), as well as refer to an ongoing event (‘if I was already in the way to your house’).

In (78b) we observe again the shift in temporal morphology transforming one hypothetical future conditional (or a Future Neutral Vivid, FNV, in Iatridou’s terms) – in (78a) – into a counterfactual conditional (or a Future Less Vivid, FLV).

---

\(^{69}\) Recall that in these contexts only *-ba* is needed to mark this *fake* Imperfective.
(78)  a. Si N\textsuperscript{70} bai bo kasa di noti, N ta txoma-u.
    if 1sg go POSS.2SG house of evening 1SG TMA call-2SG
    ‘If I go to your house in the evening, I will call you [first].’

b. Si N baba bo kaza di noti, N ta txomab bo.
    if 1sg go.TMA POSS.2SG house of evening SG TMA call.TMA 2SG
    ‘If I went to your house in the evening, I would call you [first].’

c. Si N sata baba bo kaza di noti, N ta txomab bo.
    if 1sg TMA go.TMA POSS.2SG house of evening SG TMA call.TMA 2SG
    ‘If I were going to your house in the evening, I would call you [first].’

Now let us check whether sata and sta ta have different readings even in Pres. From the sentences in (79) we get the notion of a time span as well.

(79)  a. Djon sata nxina Maria badja funana.
    Djon TMA teach Maria dance funana

b. Djon sta ta nxina Maria badja funana.
    Djon AUX TMA teach Maria dance funana
    ‘Djon has been/is teaching Maria how to dance funana.’

We know that both forms in (79) are PresPro, although from different areas of Sotavento group. But the following questions must be answered:

- do they have equivalent meanings?
- does the form in (79a) also include the meaning: ‘Djon is teaching Maria in this precise moment’?

To put it more clearly, imagine the following scenario: Pedru is waiting for Djon to join him at the coffee house; Djon is late; then Djon’s sister Lurdes enters the coffee house and Pedru, looking a little upset, asks her where Djon is / what Djon is doing. Can Lurdes’ answer be either one of the forms in (79a) or (79b)? It seems they are, in fact, equivalent; in the inland of Santiago island, the answer would be (79a) – to which, as a

\textsuperscript{70} The 1SG subject clitic is here really an M for phonological reasons – the nasal consonant is followed by a /b/, assimilating its labial feature, which also happens when it is followed by a /p/; nevertheless, it is usually written N if it does not affix to the left, cliticizing onto some other word, like dja (in which case it assumes an object-like form).

93
diagnostics, a sentence-initial gosi li ‘right now’ can be added –, whereas in Praia, for instance, it would probably be (79b).

Furthermore, and as noted earlier, sta may occur as a copula, in stage-level constructions, and sata may not, as shown in the typical contrast.

(80)  
\textit{Djon sta} / *\textit{sata duenti}.
  
‘Djon is sick.’

In (80) we see that sta may select an adjectival complement, while sata can not. There are other contexts where the two contrast, namely when there is a prepositional complement. Check this in (81a) and (81b), comparing these with the grammatical sata construction in (81c), where djongu is a verb (it refers to a brief sleep, a short nap that one takes on the couch in front of the television; it can be a noun or a verb).

(81)  
a. \textit{E sta na djongu} / *\textit{E sata na djongu}.  
  he is in nap
b. \textit{E sta na sono} / *\textit{E sata na sono}.  
  he is in sleep
c. \textit{E sata djongu} / *\textit{E sata sono}.  
  he is sleeping (a nap)

Given this important selectional distinction, let us now turn again to the contexts where either sta or sata may occur in Pres (according to the area in the Sotavento islands).

In order to clarify the non-verbal nature of sata, we may use the sentence-medial adverbial diagnostics (82b). One additional observation is that the versions with the auxiliary sta require ta marking the lower (lexical) verb – that is, the embedded verb.

(82)  
a. \textit{Mudjeris sta} *(ta) / sata ta kanta}.  
  ‘The women are singing.’
b. \textit{Mudjeris sata} / *\textit{sata tudora ta kanta}.  
  ‘The women are always singing.’
As already noted, this embedded marking is similar to perceptive constructions, where a progressive (simultaneous) reading (presumably a particular type of non-finite\textsuperscript{71}) must obtain. With the auxiliary $sta$, though, the global meaning of the sentence is different, since, besides the embedded verb, there is no other lexical predicate denoting a time with which the embedded event must coincide. In this case, any reference time must be introduced, for instance, by some adverbial expression, as in the examples seen before. Let us check the difference:

(83) a. *Kantu N txiga Maria $sta$ $ba$ $dja$(-$ba$) funana.*

‘When I arrived, Maria was dancing funana.’

b. *$N$ odja-$l$ $ta$ $dja$ funana.*

‘I saw her dancing funana.’

One more distinction between these is the following: whereas in the auxiliary construction there is only one subject (the embedded TP is defective with respect to this), with the perceptive verb an embedded subject is obligatory. As a side note, one might say it is an ECM construction, since the embedded subject assumes the position and the shape of the matrix object. Or, in a different view, we might propose that $ta$ $dja$ funana functions as a secondary predicate, with effects over the (true) matrix object. We shall return to these infinitival contexts in Chapter Six\textsuperscript{72}.

Following Baptista 2002, these constructions with $sta$ have been assumed in Pratas 2004 to have a biclausal structure. The structure description goes as follows: the matrix TP selects for a VP headed by the auxiliary; this, in turn, selects for an embedded TP,

\textsuperscript{71} See Lobo 2003 for a distinction between, on one side, the English gerund forms, called “present participles” by some authors, which enter progressive constructions and are complements of perceptive verbs (just like the ones we are analyzing here), being paraphrased in European Portuguese by either a gerundive form or a prepositional infinitive (the form $a$ + infinitive), and, on the other side, the “gerunds”, which have a different distribution and are paraphrased in European Portuguese by the infinitive. The following examples are from Lobo 2003:

\textbf{English “present participle”/EP gerundive or $a$ + infinitive} \hspace{2cm} \textbf{English “gerund” / EP infinitive}

(i) a. She was crying. \hspace{2cm} (ii) I prefer going to the beach.

\hspace{2cm} *Ela estava chorando/ a chorar. \hspace{2cm} *Prefiro indo à praia.

b. I saw the wind rustling the leaves. \hspace{2cm}

\hspace{2cm} *Eu vi o vento sacundindo as folhas.

\textsuperscript{72} Also in Chapter Six, the work in Gonçalves 1999 will be approached, with reference to the different selectional properties of auxiliaries concerning the structures they embed.
whose VP complement is headed by the lower (lexical) verb of the sentence (which stays in situ).

In English, where there is no V-to-I movement, auxiliaries are the only verb forms moving to Tense. In Capeverdean, there seems no reason to posit that the auxiliary *sta* raises to T; the adjacency between this and -ba is the sufficient condition for -ba to lower (and affix) on *sta*.

For the environments where it is a copula – in stage-level constructions (*e*/*era* participate in individual-level contexts) –, the base position of *sta* must be investigated further.

As said earlier, one of the empirical arguments to consider a biclausal structure for auxiliary constructions with *sta* – besides the fact that *sta* can bear its own TMA markers, while *sata* cannot – is the possible occurrence of an adverbial between the auxiliary form and the TMAs of the lexical verb. The assumption is, thus, that these adverbs occur between two TPs – in left adjunction to the lower TP. This position is not available with *sata*, which is a TMA morpheme – there is only one TP in this sentence.

One more interesting note at this point is that the contrast in (82b) might feed the following prediction: regarding the strict adjacency requirements in the functional domain, with *sata* (which must be adjacent to the subject and to the verb) the adverb should surface in an adjoined position, in the end of the sentence. This is a reasonable prediction, since this produces indeed a well-formed result. Check this in (84).

(84)  
Mudjeris *sata kanta tudora*.

I will abstract away from other semantic subtleties involved in this sentence, such as whether this must, or not, have a generic reading, like the English equivalent without the definite determiner (‘Women are always singing’).

Given the purposes of the present work, the relevant points have been made on the diverse meanings and implications of the previous data, in order to testify the verbal

---

73 As stated in Pratas 2002, in Capeverdean a strict adjacency between subject and verb, or between subject and preverbal TMAs when any of these are present, is required, which prevents an adverb like *tudora* from appearing left-adjointed to TP. That is why the sentence *
Mudjeris tudora sata kanta* is not considered here.
status of *sta* and the biclausal structure in which it participates as an auxiliary, as opposed to *sata*, which is a preverbal TMA morpheme. Trying to figure out if there is any possible equivalence between them, we have seen in the complexities of this sample of sentences that it depends on many different factors.

In the next point I shall address some relevant facts concerning the temporal reference construction in sentences with stative verbs as opposed to sentences with non-stative verbs. The different types of predicates, and also some specific predicates within the same group, have distinct implications with respect to TMAs. I hope it will become clear that these idiosyncrasies depend on the inherent aspectual properties of these predicates.

### 2.3.2.3 On statives versus non-statatives

From the tables (44) and (45) we know that there are some differences between verbs that denote states and verbs that do not, with respect to the selection of TMA morphemes. These distinct selectional properties depend on the aspectual idiosyncrasies of predicates, which adds one more reason to defend that temporal reference is supported by an extremely complex architecture.

As a matter of fact, even when we concentrate on non-stative predicates – having to consider the particularities of activities, achievements and accomplishments – and try to formalize a separation between the specific functions of the various functional morphemes, and between these and the verb complements, adverbials, prepositional adjuncts, discourse information and predicate semantics, that formalization seems a rather inglorious task – as I have said before. Recall, for instance, the distinct effects of *sata* and the position of *oxi* regarding verbs such as *txobi* and *txiga*.

This is also well perceived from some examples with statives. As I have said as an introductory note to tables (44) and (45), my first collection of data appeared to show some differences between *gosta* ‘like’ and *sabe* ‘know’, both considered statives: the former seemed to allow *ta*, the latter did not. In more recent eliciting meetings with my informants, however, occurrences of *ta* with *sabe* have been disclosed; they never allow for Present readings, though. The combination of *ta* with *sabe* denotes some form of posterior reading (either Future or Conditional).
In any case, the separation between statives and non-statives is a crucial tool in this matter. The main reason for this seems to be related to the following fact: the construction of temporal reference in Capeverdean gravitates largely around Aspect.

I bring to discussion again some adaptations to the prediction in Swift & Bonhemeyer 2002, on the results of combining a zero morpheme with stative verbs and with non-stative verbs: in the first case it cannot have a perfective reading; in the second case it cannot have an imperfective reading. These restrictions depend on the inherent aspectual properties of the predicates.

As noted earlier, the inherent property of the Capeverdean predicates in each group that is crucial for the reading implications of their combination with a zero morpheme is expressed in the following contrast:

i) inherent atelicity (the combination with a zero morpheme maintains the imperfective reading); 74

ii) lack of inherent atelicity 75 (the combination with a zero morpheme triggers perfectivity; the imperfective reading must be derived in a distinct combination).

The first is the case of statives; the second is the case of non-statives.

The non-stative property in ii), as also underlined in the footnote, is distinct from stating that they are inherently telic (and, consequently, it is different from stating that perfectivity is “maintained” by the combination with a zero morpheme), since activities – like faze kazas ‘build houses’ – and accomplishments – like faze si kaza ‘build his house’ – have distinct properties with respect to this (activities are atelic and accomplishments are telic), and yet they both have a perfective reading when combined with a zero morpheme.

74 The typical imperfectivity of the predicate is maintained by the absence of some morpheme that contradicts it.

75 This is different from inherent telicity.
In other words, *faze* ‘make’ is not inherently atelic, but acquires this property when combined, for instance, with a non-quantized object. This derived atelicity does not have effects on the above TMA selection.

Hence, the property in i) is reserved for the so-called states.

Now consider some practical examples. We know that the Pres of the verb *sabe* ‘know’, traditionally considered as stative, is not built with *ta*. It is marked by the bare form of the verb, or, in other words, by a zero morpheme. Its past (non-perfective) obtains with the postverbal *-ba*. This can be seen in (85).

(85)  

\begin{align*}
\text{a. } & \text{Pursor sabe } \text{ mas } \text{ di ki alunus.} \\
& \text{teacher know more of that students} \\
& \text{‘The teacher knows more than the students.’} / * \text{‘The teacher knew more…’} \\
\text{b. } & \text{N sabeba } \text{ fazeba kel bolu li } \text{ mas gosi N ka sabi mas.} \\
& \text{1SG know.TMA do.TMA DET cake LOC but now 1SG NEG know more} \\
& \text{‘I used to know how to do this cake, but now I don’t know any more.’}
\end{align*}

In certain declarative contexts, a different temporal marker may be added: *dja*. The value of *dja* can be, in some specific circumstances, an adverbial, meaning ‘already’. In other contexts, *dja + sabe* contributes an inchoative meaning, like when someone has been informed of something. The structural position of *dja*, as a temporal/aspectual morpheme, will be discussed later (in 2.3.2.4), since it reveals a special behavior with respect to subject clitics. For now, we will focus only on *ta*, *sata* and *-ba*.

As for the verb *gosta* ‘like’, likewise considered a stative, the relevant difference from *sabe* ‘know’ is not in the presence of *ta* (we shall see that the uses of this morpheme are identical among statives), but precisely in its absence. In other words, the bare form of *gosta* may have both Pres and PastPer readings – as long as the relevant temporal location is identified in the sentence –, whereas the bare form of *sabe* prohibits PastPer. We might want to compare (85a) and (86b).

(86)  

\begin{align*}
\text{a. } & \text{N (ka) } \text{ (*ta) gosta di pasia parmanhan.} \\
& \text{1SG NEG TMA like of walk morning} \\
& \text{‘I (don’t) like to take a walk in the morning.’}
\end{align*}
b. $N \ (ka)\ gosta\ di\ filmi\ onti.$

1SG NEG like of movie yesterday

‘I (didn’t) like(d) the movie yesterday.’

The negation in (86a), as we can see from the parentheses, is not determining in the prohibition of $ta$.

In (87b) we have a different example of the impossible PastPer reading for the bare verb form of $sabe$.

(87) a. $N\ (ka)\ (*ta)\ sabe\ papia\ ingles.$

1SG NEG TMA know speak English

‘I can (not) speak English.’

b. * Onti $N\ (ka)\ sabe\ kantu\ ki\ bu\ txiga.$

Yesterday 1SG NEG know when that 2SG arrive

‘Yesterday I (didn’t) know/knew when you arrived.’

We have just seen that whereas $gosta$ may occur in its bare form with a PastPer reading (86b), $sabe$ may not (87b).

Being so, we must notice that the inherent property of atelicity referred to above for stative verbs is not, for $gosta$ ‘like’, unchangeable: when it is combined with a zero morpheme but also combined with certain sentence elements (like the adverbial $onti$ ‘yesterday’ and an object that implies a closed period of time – the watching of a movie), the reading is not imperfective – that is, the tense is not Pres, as expected for the bare form of stative verbs, but PastPer.

We can, thus, add some specification to the above remarks: the imperfective reading (Pres) maintained by an inherently atelic property combined with a zero morpheme only holds for some stative verbs, like $sabe$, and also $kre$ ‘want’; for $gosta$, this only applies whenever there are no elements triggering perfective (when these elements are there, the PastPer reading is yielded).

---

76 Further investigation is needed in order to establish broader sets of these verbs.
Now let us go back to what could seem one unexpected fact: *ta* may, actually, occur with these stative verbs. The difference from non-stative contexts is that while with the latter Habitual readings (Pres and PastHab) are possible as well, with statives only this Fut reading is available (like in the consequent of conditionals, either Pres or Past).

Consider (88) and (89):

    1SG can buy DEM book LOC, Lisa TMA like of-it
    ‘I can buy this book, Lisa is going to like it.’

    if 1SG buy DEM book LOC, 3SG TMA like of-it
    ‘If I buy this book, Lisa is going to like it.’

(89) a. *N pode kumpra* kel bistidu li, *e* ka *ta* sabe.
    1SG can buy DEM dress LOC, 3SG NEG TMA know
    ‘I can buy this dress, he is not going to know (about it).’

   b. *Si N kumpra* kel bistidu li, *e* *ta* sabe.
    if 1SG buy DEM dress LOC, 3SG TMA know
    ‘If I buy this dress, he is going to know (about it).’

As said earlier, in these Future readings the distinction between statives and non-statives is neutralized. One possible reason for this is that the telic / atelic opposition is of no importance when an irrealis meaning is at stake.

We observe this neutralization in FLV conditionals as well.

    if 1SG buy.TMA DEM book LOC, Lisa TMA like.TMA of-it
    ‘If I bought this book, Lisa would like it.’

   b. *Si ningen ka flaba el, e* ka *ta* sabeba.
    if nobody NEG say.TMA 3SG 3SG NEG TMA know.TMA
    ‘If nobody told him, he would not know.’
As for *sata*, it can not mark *sabe* ‘know’, as we see in (91a). The embedded *sata* in (91b) is provided only for a contrastive effect between this and the former sentence.

(91)  
| a. Si N (*sata*) *sabeba ma bu ta benba, N ta fazeba cuscus. |
| ‘If I knew that you would come, I would have done cuscus.’ |
| b. Si N *sabeba ma bu sata benba, N ta fazeba cuscus. |
| ‘If I knew that you were coming, I would have made cuscus.’ |

We might still look at a different sentence where *sata* and *sabe* cooccur in the same sentence, but not the same clause (92a).

(92)  
| a. N *sata* kumesa ta *sabe* kuse ki e kre. |
| ‘I start to know what he wants.’ |
| b. * N *sata* *sabe* kuse ki e kre. |

In (92a) *sata* marks Progressive for the aspectual verb *kumesa*, and not for the embedded *sabe*\(^\text{77}\). The later combination is, again, barred (92b).

With *gosta* ‘like’, *sata* is possible in certain contexts, as in (93):

(93)  
| ‘I am enjoying living in Portugal.’ |
| b. N *sata* gostaba di mora Portugal, mas dipos N ten ki bai mora p’-otu kau. |
| ‘I was enjoying living in Portugal, but then I had to move elsewhere.’ |

The reasons for the difference between *sabe* and *gosta* with respect to the allowance of *sata* depend on the semantic distinction between these predicates: *sabe* is not

\(^{77}\) It could also be surprising that a Progressive marker is allowed for a verb like *kumesa* ‘start’, which in principle denotes a point in time from which a certain state of things has been settled; that is, a durative temporal marking is applied to a point. But this is a side question here.
compatible with a progressive reading, while *gosta* is. This is one more distinction to add to the previous one, with respect to the possible PastPer reading of *gosta* combined with a zero morpheme, in contrast with the impossibility of the same reading of *sabe* in identical circumstances. On the other hand, though – attesting some stative nature of *gosta* –, we have the possible Pres reading for its combination with a zero morpheme.

We cannot say, however, that for *gosta* there is an optional marking for Pres (either *ta* or the zero morpheme), since the selection of one or the other seems dependent on the temporal information available in the sentence and/or the discourse context. In other words, *ta* is dispensed with whenever a Pres reading is provided by other elements.

Bringing some more information into the debate, let us check some sentences with a different stative verb: *kre* ‘want’, which does not allow *ta* in two of the contexts that non-statives do (Pres or PastHab), neither does its bare form allow a PastPer reading.

And here we arrive, again, to the counterfactual discussion, this time around *wish* counterfactuals.

(94)  a. *N (\*ta) kre ser veterinario.*
   b. *N (\*ta) kreba ser veterinario.*
   c. *N (\*ta) kreba serba veterinario.*

In (94b), meaning ‘I wanted to be a veterinarian’, nothing is said about the actual present truth (suppose that the sentence was followed by something such as ‘and that is why I went to study in Lisbon’). As opposed to this, (94c) conveys that I am not a veterinarian. Crossing the discussed past morphology with the inherent properties (Aktionsart) of an individual-level stative predicate, the result is, as predicted, a Pres counterfactual

Recall that in Modern Greek, too, the *wish* counterfactuals are constructed with the verb ‘want’, plus a particular set of morphological markings: the matrix clause has the same morphology as the consequent of an FLV counterfactual, namely for future, imperfective and past. In Capeverdean, for things to be similar, we should have *kre* ‘want’ marked for Future and Past Imperfective.
I have shown evidence in favor of considering *ta*, in the consequents of FLV counterfactuals, the marker for Future, whereas *-ba* is the marker for both Past and Imperfective. Nevertheless, we have no *ta* marking *kre* ‘*want*’ in (94c). It does not seem problematic for the comparison to be productive, since, in Iatridou’s notations for Modern Greek, the morpheme that is identified as future in FLV counterfactual consequents is identified as a modal in the matrix of *wish* counterfactuals. Given that in Capeverdean the stative status of *kre* precludes its combination with dispensable (since redundant) TMAs in certain environments, I propose that this morpheme is also not needed and, thus, not possible in these matrix clauses. Its function might be encoded in the verb.

As for the morphology on the lower verb, we may confirm again from the contrast in (94) that the distinction between past and non-past is crucial: past participates in counterfactual wishes (conveying that the actual world is not how the subject would like it to be) – (94c); non-past participates in non-counterfactual wishes (these express a desire possible to fulfil – in this case, it is more like ‘*want*’ than ‘*wish*’) – (94a).

We conclude, thus, that the counterfactuality in *wish* contexts is not determined by ‘*want*’ morphology, but rather by the Past morphology on the embedded verb.

In (95) it becomes clear that, if we have the notion that the wish expressed can never come true – that the real world will never be as desired by the subject of *kre* –, the construction with *krebaka* + V-*ba* is preferred (95a); the option of *krebaka* + V (without the second *-ba*) becomes odd (95b). In these cases, *gosta*, in the “two *-ba*s” version, is also a good option (95c). The distinction between the two readings is also clear in (95d).

(95) a. *N kreba* serba veterinario, *mas N bai pa pursor.*

1SG want.TMA BE.TMA veterinarian but 1SG GO for teacher.

‘I wanted to have been a veterinarian but I am a teacher.’

78 Namely *ta*, in Pres, which can be due to the fact that this is an imperfective marker in these contexts, hence, not needed with a verb that is inherently atelic (imperfective).

79 The structural relation between these verbs – *kre* and the lower verb – will be discussed later.

80 In Portuguese, such as in English, this reading obtains by way of the intervention of the auxiliary *ter* ‘*have*’, in the infinitival form, selecting for the participle of the main verb (*ser* ‘*be*’, in this case):

(i) *Eu queria ter sido veterinário.*

1SG want.IMP have.INF be.PART veterinarian
b. **N kreba ser veterinário, mas N bai pa pursor.**

‘I wanted to be a veterinarian, but I am a teacher.’

c. **N (*ta) gostaba di serba veterinario, mas N bai pa pursor.**

‘I would like to have been a veterinarian, but I am a teacher.’

d. **N ta gostaba di baba Merka.**

1SG TMA like-TMA of go-TMA America

‘I would like to have gone to America.’

e. **N gostaba di bai Merka.**

‘I would like to go to America.’

In (96), the same type of relation with a non-stative lower verb shows that the morphological selection on this does not depend on this event type distinction.

(96)  

`Nha armun kreba  kumpraba karu nobu (mas dinheru ka ten).`

My brother want.TMA buy.TMA car new (but money NEG have)

‘My brother wanted to buy a new car (but there is no money).’

The bunch of sentences presented so far has provided a variety of contexts with different occurrences of *ta, -ba* and *sata*, and the differences between *sata* and *sta ta*. At this time of the relevant data description it is important to recapture some relevant questions and put some order in the basic facts and ideas.

i) should *ta* be labeled for Tense, since it marks Pres (non-stative verbs), or rather Aspect, since it participates, with *-ba*, in PastHab constructions (and Pres is, by definition, imperfective as well)?

ii) *ta* also participates in Fut and Cond (both with a Posterior property); thus, if we consider it as marking Aspect, does it have a posteriority value?

---

81 Confront this with (93b), where *sata gostaba* has a progressive meaning; in (95d) *ta gostaba* has a conditional meaning – this shows that, while *ta* and *sata* can sometimes have a somewhat similar contribution (some posteriority values), in others they have really distinct effects. The presence/absence of the Past Imperfective morphology on the embedded verb is not innocent in this distinction.

82 For the verb *gosta*, the version with *ta + -ba + V-ba* is required in counterfactuals. This proves the different implications of *ta + -ba* on *gosta*, whereas the PastHab reading of this verb is given by *-ba* alone (as we can see by the contrast between (70d) and (70e)).
iii) does -ba mark Tense, for it only occurs in contexts where some form of past is involved (although with distinct aspectual properties, and despite in counterfactuals this Tense is fake)?

iv) what to say about those environments where the lower -ba (in conditionals as well as in wish counterfactuals) brings an important semantic distinction? Can we still go on saying that this morpheme marks Tense?

v) and what about the so-called zero morpheme – does it mark Tense (present with statives and past with non-statives)? Or does it mark Aspect (imperfectivity – that is, Pres – with statives; perfectivity – that is, PastPer – with non-statives)?

vi) what about Mood? The different morphemes available in the language can participate in Indicative, Subjunctive and Conditional correspondents; hence, might we conclude, from the lack of a given mood morphology, that Mood (not modality) is simply not relevant in the language?

vii) finally, it would seem easier to put sata under the Aspect label, but besides its most common aspetual value, namely Progressive, it likewise marks Posterior in some cases. And can we be sure that, within this amalgam of features and morphemes, sata is not associated with Tense at all?

A different argument in favor of considering sata and ta as Aspect and -ba as Tense, could be that, whereas the latter can combine with any of the former, no combination between these two is possible; this could mean that these occupied the same functional head, being thus in complementary distribution. But how to apply this conclusion to complex predicates, where -ba on the embedded verb (we shall see that this is the only TMA allowed for the embedded verb) brings an obvious aspectual contribution? For the nonce we will leave these questions open.

In the next point I shall address the distribution of a different morpheme, whose status is ambiguous: sometimes it brings an adverbial contribution, sometimes its contribution is more typical of TMAs.
2.3.2.4 The puzzling distribution of dja

Among the morphemes that help build temporal reference in Capeverdean, dja is the more enthralling. A description of its distribution must include the following items:

- it may appear in the beginning of the sentence, or in the end, or both;
- whenever the subject is a DP or a free pronoun dja surfaces in a post-subject and pre-negation position;
- whenever the subject is a clitic form, dja surfaces pre-subject\(^{83}\) (which could be a sign of its peripheral syntactic position, behaving as an adverb);
- it may indeed be an adverb co-occurring with the subject clitic; not in this configuration, though, but rather in the sentence-final position.

Following Suzuki 1994, Baptista 2002 views dja “as a perfective marker (rather than a completive marker [as had been proposed in Silva 1985]), as it allows a possible continuing relevance of the action or state to the present situation.” (Baptista 2002:85)

With some predicates, like sabe ‘know’, dja actually marks the present state of things, as opposed to a past where they used to be different: Dja-N sabe means, precisely, ‘I didn’t know before but I do know now.’ This inchoative meaning also has a completive component, though – there is a completed moment in the past where this quality of knowing has been acquired.

I shall not enter into the detailed analyses of such subtleties, for the status of dja as a temporal morpheme or an adverb (this is what is really at stake for the goals of the present work) may be determined without them. The crucial fact is that for the adverbial meaning to be available, like ‘I already know’ (a possible remark when someone tells me something he/she thinks I do not know yet), we need a different construction.

In the following sentences we can check these complexities.

(97) \(N \quad ta \quad fumaba, \quad mas \ gosi \ dja-m \ para.\)

1SG  TMA smoke.TMA  but now DJA-1SG stop

‘I used to smoke, but now I have quit.’

\(^{83}\) And, thus, pre-negation ka (whereas ta and sata always appear post-negation).
In (97), *dja* contributes the concept of a certain point in a recent Past, where a different state of things has been started. The real temporal adverb in the sentence is *gosi* ‘now’.

We can easily see the different positions allowed for *dja*, with respect to the syntactic category of the subject. In (98) we have examples with a non-stative verb.

(98) a. *Dj*-e durmi. / *E dja durmi.*
   
   DJA-3SG.CLITIC sleep
   
   ‘She got asleep.’
   

In (99), we have more examples with *sabe*.

(99) a. *Dja*-N / sabe ma Clever ka ta konta-m kasi.
   
   DJA-1SG.CLITIC know COMP Clever NEG TMA tell-1SG lie
   
   ‘I know that Clever does not lie to me.’
   
   b. *Dja*-N / sabeba ma Clever ka sata kontaba mi kasi.
   
   DJA-1SG.CLITIC know-TMA COMP Clever NEG TMA tell-TMA 1SG lie
   
   ‘I knew that Clever would not lie to me.’
   
   c. *Djon (dja) sabeba ma Clever ka sata flaba kasi.*
   
   Djon DJA know-TMA COMP Clever NEG TMA tell-TMA lie
   
   ‘Djon knew that Clever would not lie.’

As said before, in some contexts with *sabe*, *dja* may mark the moment when someone gets to know – we have an inchoative reading. This does not seem to be the case in (99a) and (99b), though. The speaker is a mother, and she is talking about her 10 year old child, about whom such a declaration would be odd.

Consider the interesting contrast in (100):

(100) a. *Patricia dja sabe le.*
   
   ‘Patricia can read [already].’
   
   b. *Nha pai (???dja) sabe le.*
   
   ‘My father can read.’
Except for the case that ‘my father’ could not read until recently, (100b) with dja is odd. In (100a), assuming that Patricia is a young child, at the age of developing her reading skills, the version with dja is perfect. Hence, and from the comparison with (100) – namely with the distinction between (100a,b) and (100c) –, we could propose that postsubject dja (recall that this order is only possible with the subject being a DP or free pronoun, not a clitic) means something like ‘already’. This is not always the case, though, but I shall leave this for later.

From a simple look at some other relevant data, we may verify that dja meaning ‘already’ is, undoubtedly, the one in the sentence final position. In a preverbal position, it depends. Let us observe the sentences in (101).

(101)  a. **Dja-M bai.**
   DJA-1SG go
   ‘I’m leaving.’

   b. **Bu ta bai dja?**
   2SG TMA go DJA
   ‘Are you going/leaving already?’

   c. **Dja-bu / Dja-u bai dja?**
   DJA-2SG go dja
   ‘Are you going/leaving already?’

   d. **Dja-bu / Dja-u bai?**
   DJA-2SG go
   ‘Are you leaving?’

There seems to be a slight difference between the versions in (101b) and (101c), with the first referring more to the ‘go (somewhere)’ meaning and the latter being more common for the ‘leaving’ meaning.

---

84 Dja-u is also dja + clitic, as in Dja-bu; the difference is that the u form is traditionally taken as being the object clitic for the second person singular, while bu is the subject form. However, and as has been pointed out earlier, whenever the subject clitic has a possible phonological host on its left (it can be dja, or the preposition pa ‘for’, among other words) it sometimes assumes the object morphological form, leaning onto its left and cliticizing to that available word.
In European Portuguese we have a similar contrast between the sentence-initial and the sentence-final positions of já, with the verb ir ‘go’.

(102) a. Vais já?
   go.PRES.2SG JÁ
   ‘Are you going/leaving already?’

   b. Já vais?
   ‘Are you leaving (already)?’

Notice that both in European Portuguese and Capeverdean, the version with já/dja in the sentence-final position allows for two different meanings of the verb ‘go’: ‘to go’ (somewhere that is of the knowledge of the interventiens in the situation: kasa ‘home’, skola ‘school’, mar ‘beach’, Merka ‘America’…) and ‘to leave’. For the version with the morpheme (or only one morpheme, in the case of Capeverdean) in the sentence-initial position, only the ‘leave’ meaning is available; the place where the person is going (and which is supposedly known from the context) is not part of the relevant information.

The main distinctive feature is related to some comment on it being too early for ‘you to leave’: (101a,b) and (102a) imply that this leaving/going somewhere right now is unexpected; in (101c) and (102b) there is no such presupposition, it is an unmarked question, like when we see someone preparing to leave and ask that just to be sure, or before saying goodbye.

A simple declarative sentence like N bai ‘I go’, away from a leaving scenario, would make other people naturally ask ‘where?’, since some locative information is missing.

(103) a. N bai…85 / N ta bai… / N sata bai...
   1SG go / 1SG TMA go / 1SG TMA go
   ‘I went…’ / ‘I go /will go…’ / ‘I’m going…’

   b. Pedru bai… / Pedru ta bai… / Pedru sata bai...
   ‘Pedru went…’ / ‘Pedru goes/will go…’ / ‘Pedru is going…’

85 Recall that this is, actually, M, for phonological reasons.
Again, as seen in the distinction between *Dja-m bai* ‘I’m leaving’ and *N bai*… ‘I went…’, *dja* also seems to contribute a Tense meaning to these sentences: it eliminates the PastPer reading from *N bai*… and brings the Pres interpretation.

Crucially, if we had some other person, not intervening in the discourse situation, like in (104b) – as opposed to (101d), here repeated as (104a) –, the reading could also be about a moment not coincident with (or immediately posterior to) the utterance time, but rather a PastPer.

(104) a. **Dja-bu / Dja-u bai?**
   DJA-2SG go
   ‘Are you leaving?’

b. **Dj’-e bai? Pedru dja bai?**
   DJA-3SG go / Pedru DJA go
   ‘Has he/she left?’ / ‘Is he/she gone?’ / ‘Is Pedru gone?’

As shown earlier, in the absence of a locative complement for *bai*, *dja* changes the predicate. But, also with *dja*, if a locative complement is expressed (in Capeverdean this complement is a DP), the verb holds its meaning of ‘go somewhere’, with the correspondent temporal effects.

(105) a. **Dja-M bai.**
   ‘I’m leaving.’

b. **Dja-N sata bai kasa.**
   ‘I’m going home.’

c. **Dja-M bai kasa.**
   ‘I went home [and now I’m back].’

Interestingly, it seems that besides changing the meaning of the verb, *dja* with the bare *bai* has effects on Tense as well, since when we have a complement for *bai* the construction is Past (105c), and with the bare *bai* (the leave reading) the construction is PresPro (101a) (here repeated as (105a)). In order to have a PresPro reading for *bai* with a complement, the relevant morpheme is necessary (105b).
For (105c) we have also the information of the relative temporal proximity of this Past reading, since if the event time was for instance the day before the utterance time the sentence would be odd.

In European Portuguese *ir* ‘go’ is marked morphologically for Present and Past:

(106) a *Eu vou a casa.*
     1SG go.PRES to house

b. *Eu fui a casa (e estou de volta).*
     1SG go.PAST to house (and I’m back).’

One final (very common in spontaneous speech) example on these complex relations between *dja* and *bai*, is the one in (107), from which we may infer that whenever the locative complement is cognate – although not expressed – the verb holds the ‘go’ meaning, with all the temporal consequences for the sentence:

(107) **Dja**-M *bai dja*-M *ben.*
     ‘I went there and I’m back now.’

Now consider a different verb to be combined with *dja*. In the two sentences in (108), below, we can see more clearly that *dja* in the end is the adverbial one, while the preverbal one is more like a TMA marker.

(108) a. **(Dja)** *more tres algen.*
     ‘Three people have died.’

b. **Dja** *more tres algen dja.* / ???. *More tres algen dja?*
     ‘Three people have died so far.’

Whereas in (108a) we have a simple declarative sentence, perfectly acceptable when no specific information on possible deaths is available, the one in (108b) is only okay when, for instance, we know that there has been a big accident and at any time some news are expected on how many deaths have taken place. It could be an answer to the question in (109).
(109) Kantu ki **dja** more **dja**?
   ‘How many have died so far?’

Now consider (110), where none of the readings of **dja** – as TMA or as adverbial – would be acceptable:

   ‘Hundreds of people have died in the 46 famine.’

Again, we have evidence for these relative proximity implications of this morpheme, and that is why it is not allowed with such an event located in a distant Past.

Being more ‘die’ an unaccusative verb, indefinite subjects may stay in situ, that is, in its postverbal position. The interesting fact for the time being is that in (108b), where we have a sentence final **dja**, the initial **dja** is strongly preferred – cf. this with (101b) (**Bu ta bai dja**?), where we have the subject in its preverbal position and the initial **dja** is dispensed with. The question of whether these two facts are related – and if they are, the further question of whether this is due to structural or semantic reasons – must be investigated.

Before proceeding with a proposal for the structural position of preverbal **dja**, let me present some more contexts providing a sort of light adverbial meaning for **dja**. This is the case in (111), where it occurs to the left of the clitic and to the right of a DP.

(111) **Djon, dji’-e sabe ma-N ka ta gosta d’-es filmi-li.**
   ‘Djon, he knows that I will not like this movie.’

We have a different, more subtle, example of this in (112a), as opposed to (112b).

(112) a. **Djon dja txiga, gosi nu pode kume.**
   ‘Djon has (finally/just) arrived, now we can eat.’
   b. **Bu sata dura ku ben mas Djon dja txiga dja.**
   ‘You are late, but Djon has already arrived.’
Some sort of light adverbial interpretation is also available for it with the weather verb \textit{txobe} ‘rain’ marked for PresPro (113):

\begin{align*}(113) \quad \text{a. (Dja) sata txobe.} \\
&\text{‘It’s raining.’} \\
\text{b. (Dja) sata txobe } \text{dja.} \\
&\text{‘It is raining already.’} \\
\text{c. * Sata txobe } \text{dja?/ Dja sata txobe } \text{dja?}
\end{align*}

In contrast with \textit{more}, however, in (113b) we observe that with \textit{txobe} the final \textit{dja} may occur without the initial \textit{dja} (although it is more common with it). The singular final \textit{dja} is only truly disallowed in the question in (113c). I propose that these particular properties of morphemes participating in the temporal reference also depend on specific semantic and syntactic properties of predicates (notice that whereas for weather verbs there is no subject, for unaccusatives there is a subject, which may surface in a postverbal position whenever it is indefinite).

Let us now take a closer look at the possible combinations of \textit{dja} with other TMA morphemes. So far, we have seen that preverbal \textit{dja} (pre-subject clitic and/or post-subject DP) may combine with the sentence-final adverbial \textit{dja} and also with \textit{sata}.

In the examples in (114) we confirm that it can combine with \textit{sata} (114a), but not with \textit{ta} in the non-negation context in (114b). The combination with \textit{ka} plus \textit{ta}, however, is possible (114c).

\begin{align*}(114) \quad \text{a. Dja-}\text{-N sata bai.} \\
&\text{DJA-1SG TMA go} \\
&\text{‘I’m going.’ / ‘I’m leaving.’} \\
\text{b. * Dja-}\text{-N ta bai.} \\
\text{c. Era pa-}\text{-m bai skola, mas gosi (dja)-N ka ta bai mas.} \\
&\text{‘I was supposed to go to school, but now I won’t go anymore.’}
\end{align*}
The reason for the ill-formedeness of (114b) cannot be strictly structural (following some line of reasoning on there being only one preverbal slot for TMAs\textsuperscript{86}, which would thus be in complementary distribution).

The grammatical sentence in (114a) shows that the prohibition of $dja + ta$ is not due to any structural reason – there are two slots available for $dja + sata$. Hence, the prohibition of $dja + ta$ must be due to semantic reasons, which is perfectly acceptable: $ta$ and $dja$ bring incompatible temporal information to the sentence. This is confirmed by (114c), where the negation has a crucial role.

One possible suggestion is that the scope of $dja$ operates on the negation marker $ka$ and not on $ta$; somehow, the intervention of the negation eliminates the incompatibility between the semantic/aspectual contributions of $dja$ and $ta$. A further observation is that this preverbal $dja$ in (114c) operates together with the negation $ka$ and the final $mas$ (in a fashion similar to the relation established between the preverbal and final $dja$'s), which allows for different combinations.

In any case, the complex contribution of preverbal $dja$ seems to allow for a completive reading, since “it allows a possible continuing relevance of the action or state to the present situation” (recovering the interpretation proposed in Baptista 2002). We may apply this to (114c): in this context $ta$ marks Future, and the negation of this eventual future event (as opposed to the future event itself) is compatible with $dja$. The sentence conveys that $gosi$ ‘now’ a certain state of things has been established such that onwards I will not go to school. In other words, we have the reference to a close moment in the Past where the progressive action has started; $dja$ marks the left boundary of the time span to which the predicate applies.

Hence, preverbal $dja$ cannot be a strict PastPer marker whatsoever; PastPer markers (such as a zero morpheme with non-stative verbs) may combine with Past references that need not be close; moreover, they are independent of the current state of things. We might say that, instead, the semantic contributions of $dja$ could be approached with some of the analysis proposed for the complex English Perfect.

Let us now take a look at some more combinations of $dja$.

\textsuperscript{86} There is one additional complication to this, though: for the morphemes $-ba$, $-du$ and $-da$ (the last two for passives), which are generated pre-verbally and affix post-verbally by lowering, to be excluded from this hypothetical competition (and we know they are, since they may co-occur with $ta$ and $sata$), this should take place in the morphological component, not in the syntactic one.
Also the verbal morpheme *ka* ‘finish’ may combine with preverbal *dja*, as in (115), ahead. This *ka* is almost homophonic with the negation. In fact, though, the vowel of this morpheme can safely be transcribed as /a/, while the vowel in the negation *ka* is more like a schwa.

This *ka* ‘finished’ is not very common in other Cape Verde regions; together with the TMA *dja* it has, actually, a completive reading: *Juau dja ka fazi TPC* ‘Juau has (just) done all the homework’ Recall from a previous footnote (ft 39, on page 64) that the temporal/aspectual verb *kaba* ‘finish’ is not *ka+-ba* (Past Imperfective), but a full verb form (whose bare occurrence stands for PastPer; the Past Imperfective is *kababa*). It may be derived from the Portuguese *acabar* ‘finish’.

It is possible that the phonological shape of *kaba* has been reanalyzed by the speakers, who suppressed the final syllable as if it were a mark for Past Imperfective, giving way to the Perfective form *ka*.

   Eder / dja-3SG finished get.dirty
   ‘Eder, he just got all dirty.’

b. *(Dja)-)*N  *ka*  kume.
   ‘I just finished eating.’

c. *(Dja)-)*N  *ka*  kumeba.
   ‘I had just finished eating.’

The verbal nature of this *ka* is confirmed by the fact that it may bear TMAs itself, like in (116) (notice that the negation *ka* always appears pre-TMA morphemes *ta* and *sata*).

(116) a. *Manhan, N* ta  *ka*  sabe, N ta konta-u.
   ‘Tomorrow I will know the rest of it and I will tell you.’

As for the negation *ka* (117):

(117)  *Ka dja*-N  kume-l *dja*?
   NEG DJA-1SG eat-3SG DJA
   ‘Haven’t I eaten that already?’
As for the combination between preverbal *dja* and *-ba*, it is productive in contexts clearly pointing to a time anterior to the reference time (118a). We observe, however, that only a very specific scenario allows for a final *dja* to be completely expected – confront (118b) with (119).

(118) Question:

*Modi ki bu sabeba ma Paris e sabi?*

‘How did you know that Paris is lovely?’

Answer:

a. (Pamodi) *Dja-*m *baba* la.

‘(Because) I had gone/been there [before].’

b. ???. *Dja-*m *baba* la *dja*.

(119) Question:

*Kel la era purmeru bes ki bu baba Paris?*

‘Was that the first time you went to Paris?’

Answer:

*Nau. *Dja-*m *baba* la *dja*.

‘No. I had been there already.’

Thus, what we have in (118a) is a sort of a Pluperfect reading, which is built by *dja* + V-*ba*. This temporal specification – previous to the Past reference time –, however, does not necessarily imply an adverbial input. Unless, obviously, the final *dja* is added – and this, as we see in (118b) and (119), is only welcome in certain discourse contexts.

Summarizing some ideas: so far, we have seen that preverbal *dja* changes the meaning of the verb *bai* from ‘go somewhere’ to ‘leave’ whenever a locative complement is not expressed (or given, by some discourse information); in this case, with the bare verb form *bai*, it also changes the Tense from past (PastPer) to present (Pres); we have also seen that some sort of light adverbial interpretation is available to it with unaccusative predicates like *more* ‘die’, or weather verbs, like *txobi* ‘rain’.

There is more to *dja*, though. One of the puzzles still to be answered is the fact of it obligatorily surfacing on a pre-subject position whenever this one is a clitic form. We
shall address this later. For the time being, it is my belief that *dja* might provide enough questions and possible analyses for an entire dissertation.

As it is not the main character of the present work, for reasons of space and time the next examples illustrating some more subtleties of *dja* will have the respective notations and paraphrases, but no further commentaries will be added. In some cases, I confront these with other sentences without *dja*, only to make the effects of its presence clearer.

(120) a. *N ka sabe.*
   ‘I don’t know’

   ‘I know.’

c. *Dja-N sabe dja.*
   ‘I know it already.’ [comment to someone who is telling me something]

d. *Dja-N sabeba.*
   ‘I knew.’

e. *Dja-N sabeba dja.*
   ‘I knew that before.’

f. *Dja-N ka sabe mas. / N ka sabe mas.*
   ‘[I used to know but] I don’t know anymore.’

g. *Dja-N ka sabe.*
   ‘I’ve just been told about it.’

The distinction between the *ka* in (120f) and the *ka* in (120g), even if it does not become clear from the different vowel (an /ə/ in the first and an /a/ in the second), is given by the obligatory *mas* in the end, which has the function of getting this notion of some fact being negated from a certain point on. Here, thus, *dja* and *mas* operate together, as an adverbial form, much like both *dja* and *dja* in (120c) and (120d). Recall what has been proposed concerning the same combination (114c).

Now consider the following examples:
(121)  

a. **Dja dura ki e mori.**  

DJA take.long that 3SG die  

‘He died a long (time) ago.’

b. **Dj’-e mori tempu.**  

DJA-3SG die time  

‘He died a long (time) ago.’

(122)  

a. **Es ta dura ku ben.**  

3PL TMA take.long with come  

‘They will take a long time to come/arrive.’

b. **Es dura ku ben.**  

3PL take.long with come  

‘They took a long time to come/arrive [but they are here now].’

c. **Es sata dura ku ben.**  

3PL TMA take.long with come  

‘They are taking a long time to come/arrive.’

d. **Es staba ta dura ku ben. / Es sata duraba ku ben.**  

3PL AUX.TMA TMA take.long with come.  

‘They were taking a long time to come.’

e. **Dja dura ki es ben.**  

DJA take.long that 3PL come  

‘They came/arrived a long ago.’

Compare especially (122b) with (122e): in both, the arriving time is Past, but in the former this arriving moment marks the end of the time span which *dura* refers to, whereas in the latter the arriving moment marks the beginning of the time span.

(123)  

a. **N ta bai y N ta ben.**  

1SG TMA go and 1SG TMA come  

‘I will go and come quickly.’ / ‘I will go but I won’t take long’

b. **Dja-N ben.**  

‘I have arrived.’ / ‘I’m here (already).’

---

87 The matrix clause has no subject; the only (and obligatory) referential subject in the sentence is in the embedded clause.
A. Question where the speaker does not know whether the fact took place or not.

(*) Bu sabe si Patricia dja le livru?/si dj’-e le livru?/* si e dja le livru?
   ‘Do you know if Patricia read the book?’

B. Question where the speaker knows that the fact took place, and wants to know whether the listener is aware of it or not.

*(djaj)-u sabe ma Patricia dja le livru?/ma dj’-e le livru?/*m’-e dja le livru?
   ‘Do you know that Patricia read the book?’

(125) a. Dj-e le es livru li.
   ‘He read this book.’

b. Dj-e le livru dja.
   ‘He read the book already.’ (some judgment about having read it too fast)

c. * E le livru dja.

d. * E le livru dja?
   (this question with no initial dja is much worse than ???More tres algen dja?)

In the following list of sentences, examples b. and e. describe a different relation between a time period and an accomplishment (faze si kasa ‘build his house’) are provided. In a., the relation is between a period of time and an activity.

(126) a. Djon ten tres anu na faze kasa, mas gosi el e piscador.
   ‘Djon spent three years building houses, but now he is a fisherman.’

b. Djon ten tres anu ta faze si kasa.
   ‘Djon spent three years building his house.’

c. Dja ten tres anu ki Djon faze si kasa. / Djon, dja ten tres anu ki e faze si kasa.
   ‘Djon finished building his house three years ago.’

d. Djon sata faze kasa dja ten tres anu n’el.
   ‘Djon is building [his] house has three years in it.’

e. Djon dja ten tres anu ta faze si kasa. / Dj’-e ten tres anu ta faze si kasa.
   ‘Djon / he has three years (now) in building his house.’

In (126b) we have a single house; this concrete object triggers telicity (an accomplishment), while the predicate in (126a), with a plural non-concrete DP in the
object position, is atelic (an activity). This has effects on the morpheme selected – na or ta – to introduce the (apparently) non-finite verb faze.

The consequences of the Aktionsart of predicates on their selectional properties with respect to TMAs have been addressed earlier. These examples illustrate more possibilities for dja, in this context of time duration. This durative period is the one of the activity itself (126d,e) or the one that mediates between the accomplishment and the reference time (since ten ‘have’ – in the existential sense – is Pres, the reference time coincides with the utterance time).

2.3.2.5 Some structural clues regarding dja

I have hitherto illustrated some relevant observations on the status of dja. It can be an adverbial in certain contexts (clearly in the sentence-final position, where it may cooperate with a sentence-initial dja of a different nature; sometimes in a preverbal position\(^88\)). These adverbial contexts are not problematic, though. The most puzzling combination is the pre-subject clitic dja when it does not have an adverbial contribution, but is rather a TMA.

My first temptation has been to propose that the initial dja is obligatory inasmuch as some material is needed in Spec,TP (to satisfy the EPP\(^89\)). This dja could, then, be analyzed as an argumental word with an aspectual meaning, of the type proposed in Costa 2004, or in a construction similar to English locative inversion, like ‘In the room came John’ (see Bissel-Doggett 2004 for a discussion on English locative inversion).

Also in Holmberg 2005, on the evidence from Finnish on little pro, it is observed that the fronting of a temporal or locative adjunct satisfies the Finnish version

---

\(^88\) This seems to violate adjacency requirements, but this “violation” is only possible given the light morphological weight of dja – which is important, since the adjacency requirements operate at the morphological level.

\(^89\) The Extended Projection Principle (which will be addressed in greater detail in 5.3.2) roughly determines – in spite of all its various versions over the last decades – that the subject position be filled. Although this is a position typically reserved for subjects, and in its first formulations the EPP required that all sentences have subjects, as we shall see, with time two transformations occurred:

i) even when sentences have subjects and they are clearly not in this position, something must be there only to satisfy the EPP;

ii) this ‘something’, thus, does not have to be a potential subject; it can be a null category (like pro – even in cases where the subject is not null) or, for instance, a temporal marker.
of the EPP. More generally, the author states that certain adverbials that are referential (locative, temporal, instrumental – but not manner or reason) can check the EPP.

As a matter of fact, this resort would prove not to be necessary for Capeverdean. The first and main reason for rejecting *dja* as surfacing on Spec,TP, given the ultimate goals of the present work, is that the positing of this position for a temporal marker has only one purpose: satisfy the EPP.

In Chapters Five and Six it will become clear why this is not such a great achievement for Capeverdean *dja*:

i) the EPP may be subject to parametric variation;

ii) it may have a negative value in Capeverdean (any syntactic operation that only serves to satisfy the EPP is, thus, not necessary; under minimalist concerns, this means it is not allowed).

There are two reasons for this.

i) nothing seems to be needed in the Spec,TP position when the subject of the sentence either does not have any semantic content (expletive constructions), or does not need to move from its base position (indefinite subjects of unaccusative predicates) – I shall propose that, in both these cases, Spec,TP is not projected.

ii) neither the non-adverbial nor the adverbial *dja* “need” this position – the first one surfaces in one of the adjoined functional heads, labeled T, as I proposed for TMAs; the latter surfaces as an adverbial adjunct, adjoined to VP.

### 2.3.2.6 Concluding remarks on TMAs

Over the last pages I have extensively described the distribution of TMAs: preverbal *ta, sata* and postverbal *-ba*, and also preverbal (and pre-subject clitic) *dja*.

I have presented some consequences of the stative versus non-stative status of predicates on TMA selection and some crossings of grammatical Aspect and Aktionsart concerning temporal reference construction.
The main conclusion is that it would be too arduous a task to determine whether a given morpheme should surface on Tense, whereas other would surface on Aspect or some differently labeled functional head, since their specific contributions depend on various, intertwined factors. In other words, temporal reference is, in Capeverdean, typically compositional.

I will try to account for this with my proposal on the syntactic structure for the Capeverdean functional domain, in section 2.4. I will argue that there is no evidence for a strict correspondence between each of these morphemes and a specific functional head such as Asp, Mood or any other of this sort. All of them cooperate, in ways that are marked by various interdependencies, in temporal reference reference construction. Hence, they may all be labeled T.

In the next subsection, some data will illustrate the distribution of the copula \( e/era \) (Present/Past forms) and the negation marker \( ka \); which co-occurs with words like the adverbial \( nunka \) ‘never’ and the quantized operator \( ningen \) ‘nobody’.

### 2.3.3 Negation: \( ka \) and the copula \( e/era \)

Following the complex contributions of the small set of Capeverdean TMAs, we must address the questions around the syntactic position of the negation \( ka \).

First of all, we must find out what the categorial status of \( ka \) is. The direct conclusion from the available diagnostics is:

- \( ka \) is a head.

The lines of argumentation go as follows.

It can not be focalized:

(127)  \( Ami, N \) gosta txeu di kaxupa, mas abo, bu *\( ka / nau \).  
1SG, 1SG like much of kaxupa but 2SG, 2SG NEG / no  
‘I like kaxupa a lot, but you, you don’t.’
It cannot occur isolated:

(128) Question:  
\[Bu ta \quad ben \quad ku \quad mi?\]  
\[2SG \quad TMA \quad come \quad with \quad 1SG?\]  
‘Do you come /are you coming with me?’

Answer:  
* Ka / Nau.

No.

One side information on the distribution of \(ka\) is the following: it co-occurs with negative DP’s or other negative operators, keeping the negative value of the sentence (that is, there is negative concord in Capeverdean, under a spec-head configuration):

(129) a. \( \text{Ningen} \quad *(ka) \quad \text{gosta di mi.}\)  
\[\text{NEG} \quad \text{NEG} \quad \text{like} \quad \text{of} \quad 1SG\]  
‘Nobody likes me.’

b. \( \text{Tioxi N} \quad *(ka) \quad \text{bai Portugal}.\)  
\[\text{NEG} \quad 1SG \quad \text{NEG} \quad \text{go} \quad \text{Portugal}\]  
‘I have never gone/been to Portugal.’

Contrast with the correspondent sentences in European Portuguese:  

\[\text{\textit{Tioxi} seems derived from the Portuguese até hoje ‘until today’, which occurs in negative or positive constructions. In Capeverdean, however, it only co-occurs with \(ka\), with the value of ‘never until today’}.\]

\[\text{\textit{Only as a side note, a projection NegP for European Portuguese is defended in Gonçalves 1999, following the proposal in Zanuttini 1996 for the Italian sentential negation non. NegP depends on the presence of a functional category which can satisfy the selectional requirements of its head: this functional category is T. Gonçalves 1999 departs from this proposal in order to account for certain restrictions concerning negation in complex predicate contexts, at least with control and raising modal verbs (such as querer ‘want’ and poder ‘can’). The author proposes a new hypothesis for Zanuttini 1996’s formulation, adding the need for T to be active in order to license NegP. Consider the following sentences (adapted and translated from Gonçalves 1999:289):}\]

(i) a. \( A \ \text{Ana quis-lhes} \quad (*não) \quad \text{contar a verdade.}\)  
\[\text{the} \quad \text{Ana} \quad \text{wanted-3PL.CLITIC} \quad \text{NEG} \quad \text{tell the truth}\]  
‘Ana wanted to tell them the truth.’

b. \( A \ \text{Ana pode-o} \quad (*não) \quad \text{ter encontrado}.\)  
\[\text{the} \quad \text{Ana} \quad \text{can-3SG.CLITIC} \quad \text{NEG} \quad \text{have found}\]  
‘It is possible that Ana didn’t find him.’

\[\text{Her new formulation goes as follows: Neg (and, thus, NegP) can only occur in the embedded domain when the embedded T is not defective. The underlying reasoning is the following. The author has assumed that there is an embedded TP in the above sentences. We know that this embedded T is defective}\]
(130)  a. _Ninguém (*não) gosta de mim._
      nobody not like of me

    b. _Eu nunca (*não) fui a Portugal._
      1sg never not went to Portugal

Now that we have confirmed the head status of _ka_, the following step is to determine on which functional head it surfaces. Consider (131):

(131)  a. _Wosvaldu ka era riku. / * Wosvaldu era ka riku._
      ‘Wosvaldu was not rich.’

    b. _* Wosvaldu ka e riku. / Wosvaldu e ka riku._
      ‘Wosvaldu is not rich.’

    c. _Wosvaldu ka riku._
      ‘Wosvaldu is not rich.’ / * ‘Wosvaldu was not rich.’

    d. _* Wosvaldu riku._

Notice, now, the interesting phenomena:

  i) the present copula _e_ is in a pre-negation position, contrasting with all other verb forms (131b);

  ii) in the presence of _ka_ the present copula is preferably null (131c);

  iii) a null present copula is not allowed in the non-negation context (131d);

  iv) as seen in the second paraphrase for (131c), copula ‘deletion’ is ruled out with the phonologically heavier past copula _era_; it is ruled out also with the much less productive perfective form _foi_, which does not trade positions with the negation either (E (*foi) _ka_ *(foi) _simpatiku ku gentis ki txiga_ ‘He was not nice with the people who came’).

My proposal will be that this is due to the relative position of _e/era_ and _ka_: the copula form and the negative marker appear in adjunction, under the label T.

---

since there is complex predicate formation: according to the author’s analysis, in complex predicates there is raising of the embedded verb to the matrix domain. This can be rendered visible by way of the clitic climbing – one of the traditional diagnostics for complex predicate detection. If the embedded T were active, it could check the features of the embedded verb, which, thus, would not need raise (this is to say that it would not be allowed to raise). In this case the embedded sentential negation is available and we would have some well-formed sentence like: _A Ana quis não contar a verdade aos pais._
That the negative marker *ka* is an adjunct to T is not surprising at all (in a view different from Gonçalves 1999, this position for the negative marker in European Portuguese has been considered in Matos 1999). As for the copula, considering that it bears no meaning other than temporal information, this T adjunction position is not surprising as well. Furthermore, this configuration accounts for the two facts above.

The line of reasoning goes as follows: these facts are the result of morphological operations; these operations can take place inasmuch as, under this head adjunction configuration, the required conditions are met.

Let us now look at the details:

i) when the copula is in the present form, *e*, its morpho-phonological weight determines that it switches positions with *ka*; since the past copula *era* is morpho-phonologically heavier, this form stays in its post-negation position;

ii) as for the null Present copula, the fact that it is disallowed in non-negation contexts favors the hypothesis of an (optional) morphological merger\(^{92}\) between *e* and *ka*; this is possible indeed, since the adjacency conditions required for this operation are met in this configuration.

The fact that the morphological operation in i) is obligatory (there is no *ka e* order) and the operation in ii) is optional (although the null copula is preferred\(^{93}\)) depends on other

---

\(^{92}\) This operation has been first proposed in Marantz 1984: a relation between X and Y may be replaced by (expressed by) the affixation of the lexical head of X to the lexical head of Y.

\(^{93}\) In order to test whether the present copula plus negation are allowed (that is, whether copula deletion in negative contexts is not obligatory), we must test with person+number combinations other than 3sg. This is so because of the phonological similarity between the 3sg subject clitic and the present copula (see Itzel 2006 for a historical overview of the two forms, concluding that – as opposed to what has been argued elsewhere – the copula and the 3sg clitic have different origins and natures). Consider the following:

(i)  
   a. *Wosvaldu e ka riku.*  
   ‘Wosvaldu is not rich.’
   b. *N / Mi e ka riku.*  
   ‘I am not rich.’
   c. *Bu / Bo e ka riku.*  
   ‘You are not rich.’

Without any further information, one could suppose that in (ia) we have a null copula and a clitic doubling context: DP + clitic. However, if we check (ib) and (ic), we observe that the subject clitic is not
independent factors that distinguish word orders based on weight reasons and the merger of two morphemes (and I shall not go into its details here).

When there is no visible copula, the precise temporal reading is (in a fashion similar to what we have seen for other morphemes in the previous section) dependent on other pieces of information available in the sentence. Take (132a), where we have a passive construction (in which the copula is not allowed). Although the fact we are referring to is located in PastPer, its consequences extend in time, involving the notion of Pres (akin to some English Perfect values – check both paraphrases for (132a)): the passive morpheme \textit{-du} is, thus, considered to mark Pres, while the much less common \textit{-da} marks Past, as in the impersonal passive in (132b).

(132) a. \textit{Djon (*e) tradu di trabadju.}  
\begin{flushright}
‘Djon has been fired.’ / ‘Djon is unemployed.’
\end{flushright}

b. \textit{Kelbes ta badjada txeu (mas gosi ka sata badjadu mas).}’  
\begin{flushright}
‘In the old days people danced a lot (but now they are no longer dancing).’
\end{flushright}

For the proposal of merger with the present copula to make sense, the preverbal (and also pre TMA, as we have seen before) negation marker \textit{ka} must: i) be a head; ii) surface on a T position. I hope to have shown that no one of these conditions is a problem.

Given all these arguments, I maintain that every piece of functional morphology in Capeverdean sentences, including Negation marker \textit{ka}, is generated on a functional head under the label T. The only morpheme that does not surface in this very position is \textit{-ba}, which is generated there and affixes to the right of the verb (in Chapter Three I will analyze this phenomenon in detail).

To be more specific, the TMAs \textit{ta} and \textit{sata}, and also \textit{dja}, the copulas \textit{e} and \textit{era}, and finally the negation marker \textit{ka}, surface on successively adjoined T heads.

---

allowed in these contexts. In Pratas 2002 I have shown that the X° forms of the pronominals are disallowed with the copula; hence, at least a free form is obligatory (this fact led me to consider that free forms behave in some cases as X°, having an ambiguous status, but this particularity is not relevant here).

The important fact is that in (ia) the form \textit{e} cannot, thus, be a 3sg clitic. Hence, it must be the lexicalized Present copula, which, obviously, is allowed in negation contexts (although preferably null).
There are two lines of reasoning leading to (and converging in) this proposal. One of them regards the semantic contribution of these morphemes; the other depends on positional arguments. As follows:

i) each of these morphemes is, in one way or another, engaged in the temporal reference construction; they may do it separately, but also by combining with each other and with other semantic and discourse information available;

ii) some pieces of structural evidence, like the possible deletion of the Present copula *e* in the presence of the negation *ka* and, as we will see in the next section, also the fact that the subject clitic appears post-*dja*, favor the head adjunction position for them all.

In the next section I will address some other structural arguments against Capeverdean TMA markers instantiating different heads in the IP domain. Recall that the main argument against an array of functional projections comes from the fact that TMA morphemes are string-adjacent. An analysis according to the proposal in Cinque 1999 would predict the possibility for XPs to occur in between the several heads, occupying their specifier positions. In section 2.2 and in the present section we have seen that this prediction does not hold in Capeverdean.

Taking minimalist concerns (of economy and explanatory adequacy) into consideration, and searching for support in previous proposals (such as Bobaljik & Thráinsson 1998), I will contend that Capeverdean has a minimal functional architecture. All morphemes that surface in this domain may be accommodated by one maximal projection alone: TP.
In the previous sections we have seen that in Capeverdean any intervention of adverbs or any other XP in between TMA markers, or between TMA markers and the verb, is disallowed. This favors the existence of adjacency requirements hard to explain under an exploded functional domain.

Examples, among many, illustrating these word order restrictions are the ones in (133) and (134).

(133)  a. * Ana Maria ka dretu ta gia karu. (manner between functional morphemes)
        Ana Maria NEG right TMA drive car
       b. Ana Maria ka ta gia karu dretu.
          ‘Ana Maria is not a good driver.’

(134)  a. * Nastasi ta oxi kume na kuzinha. (temporal between TMA-verb)
        Nastasi TMA today eat in kitchen
       b. Oxi Nastasi ta kume na kuzinha.
          ‘Today Nastasi eats in the kitchen.’

In (135) we can see the contrast between the corresponding representations:

(135)  a. TP
       Nastasi
            T’
                 T
        T
        ta
        Adv
        VP
        * oxi
        V
        PP
        kume
        na kuzinha
We have also seen that there is no positive evidence to claim that each TMA marker occupies a separate head position.

On the contrary, there is evidence to claim that they surface in head adjunction under the T label. This evidence is twofold:

i) Negative evidence: it would be difficult to dispose these TMA markers in distinct functional projections, since the contribution that each of them may bring to the sentence depends on the semantic properties of predicates, on other functional morphemes, adverbial expressions and discourse information;

ii) Positive evidence: their obligatory adjacency is well attested; all of them conjoin to build temporal reference; even the negation marker *ka*, which has also been shown to be a head, must surface in the temporal domain, as we can see by its possible merger with the Present copula *e*.

In this whole section I will assume a Split,TP parameter (Bobaljik 1995; Thráinsson 1996; Bobaljik & Thráinsson 1998). According to these proposals, the inventory of functional projections is subject to parametric variation.

In subsection 2.4.1 there is a short summary of the considerations involved in the positing of a Split,IP parameter, considering Capeverdean case. In 2.4.2 it is shown how, given all arguments, Capeverdean TP may be an almighty functional projection.
2.4.1 The Split ,TP parameter and the position of subject clitics

Capeverdean shows no evidence for splitting I into T and Agr. Assuming, as I am, that an Agreement node is a point of parametric variation\(^{94}\), one is led to the conclusion that there is no evidence for such node in the language.

As mentioned briefly in section 2.1, there are two possible lines of arguments with respect to this:

i) the language displays no overt verbal agreement morphology (the verbal paradigm can be illustrated with *badja* ‘dance’: *N* badja ‘I danced’; *bu* badja ‘you danced’; *e* badja ‘he/she danced’; *nu* badja ‘we danced’; *nhos* badja ‘you danced’; *es* badja ‘they danced’); hence, no strong agreement features trigger verb movement to such a functional position;

ii) no positional properties (that I am hitherto aware of) favor such a projection, since, as opposed to French, for instance, there is no need for two different heads as landing sites for any verb unit\(^{95}\).

Baptista 2002, however, contends that Capeverdean subject clitics are heads in Agr (Baptista 2002:250).

Before proceeding with my argumentation against this view, it is necessary to take a look at the available pronominal forms in the language.

---

\(^{94}\) In a nutshell, the proposal in Bobaljik & Thráinsson 1998 is that languages vary with respect to the functional projections: the IP can be split into an agreement projection and a tense projection in some languages and be unsplit (or fused) in others. Hence, functional projections are not universally instantiated, which means that the child needs evidence for their presence in the language that she is acquiring. This evidence may be morphological (clearly distinct tense and agreement markers) or strictly syntactic (in other words, even in the absence of such clear morphological distinctions, there may be V-to-I movement in the language, for instance, depending on some feature checking factors.

\(^{95}\) For Capeverdean, where the auxiliary *sta* also appears to the right of the negation *ka*, and may appear to the left of certain adverbs, it has been proposed in Baptista 2002 that there is a biclausal structure – a matrix TP and an embedded TP. In other words, in these environments the auxiliary selects for a TP complement. I have been following this proposal.
I have assumed earlier (Pratas 2002) that Capeverdean lexicalizes pronominals of the three types predicted in Cardinaletti & Starke 1994\textsuperscript{97}: strong forms (XP) (they are called emphatic in (136)), and also both deficient forms: non-clitics (with a reduced structure compared to proper maximal projections; these are called free forms in (136)) and clitics (X°). Free forms – non-clitics – have an ambiguous behavior (XP or X°). In sum, the table in (136) includes my recent adjustments of the labels in each group.

Resuming the view on Capeverdean subject clitics as heads in Agr (Baptista 2002:250), this proposal seems a resort to license referential pro in the rare environments where this null referential subject is, apparently, allowed. The argumentation goes as follows: the 3SG clitic e in Agr would be endowed with third person features (the features of the null element el), leaving the argumental subject position, Spec,AgrSP, free to be filled with a pro (Baptista 2002:256)).

One possible empirical reason to consider a referential pro in Capeverdean matrix clauses is, as the author defends, the sentence in (137). In my previous works it

\textsuperscript{96} Curiously, a form homophonous with 3sg object clitic occurs in some possessive environments. This form leans onto words other than verbs, like a noun – as kasa-l nha dona ‘house of my grandmother’ – or prepositions – as dentu-l kasa ‘inside of house’. Crucially, this possessive -l has not the same effect as the object clitics on the stress shift of the phonological word obtained. See 4.1 for the description and analysis of this shift triggered by the affixation of object clitics (this is a phenomenon sensitive to the categories of morphemes involved – in other words, it is truly a morpho-phonological phenomenon).

\textsuperscript{97} A similar proposal has been presented in Baptista 2002, although with differences that I shall not discuss here.
has been shown that, taking every other empirical evidence in the language, the null subject here is expletive, not referential, as we can see by the paraphrase in English:

(137)  
\[ E n h a \ p a i . \]

is my father

‘It’s my father.’

In the present work, I maintain that Capeverdean does not allow null referential subjects in matrix clauses (this will be discussed in detail in the next chapters). I also assume that no expletive pro seems to exist either: in the so-called expletive constructions no subject position is projected (this proposal will also be developed in Chapter Five). Under this light, no resort is needed to leave the Spec,AgrSP position free for referential pro.

As for the concrete proposal that subject clitics are the spell out of Agr features on Infl, I contend, against this view, that they are not mere bundles of agreement features (inflectional morphemes generated on the finite verb in the lexicon).

They actually display various properties of argumental clitics (syntactic entities), as we shall analyze below. Depending on the environment, they may cliticize phonologically onto their left: the host may be the TMA \( dja \), the complementizer \( ma \) or the preposition \( pa \), for instance. In some of these left-leaning contexts they may have the morphological form that is assumed to be the object form (this “object” forms could be taken as being morphologically marked for Accusative/Dative case). In certain cases, there are phonological consequences of this on the host as well, namely a vowel deletion, like in \( dja-e\ txiga \) ‘he arrived’.

In (138) we have instances of this phonological leaning of the clitic onto the temporal marker \( dja \) (138a), the complementizer \( ma \) (138b) and the preposition \( pa \) (138c):

(138)  
\[ a. \ D j a \ b u / D j a-u \ s a b e \ m a \ D j o n t a \ k o n t a k a s i t u d u \ d i a . \]

TMA-2SG         know COMP Djon TMA tell lies all day

‘You know John lies all the time.’

---

98 Although the two forms – the one with the subject form, \( bu \), and the one with the object form, \(-u\) – are typically equivalent, the first one is more common in interrogative contexts and the second one is more common in declarative contexts.
b. Dja-N sabe ma-u ta bai undi-m.
dja-1SG know COMP-2SG TMA go where-1SG
‘I know that you will visit me.’

c. E pa-u bai undi-m.
BE PREP-2SG go where-1SG
‘It’s for you to visit me.’

The shape of the resulting phonological words above is, depending on the areas and even on the speakers, something like /djo/ , /mo/ and /po/, respectively.

In (138c) we observe counter-evidence to the generalization “prepositions cannot support clitics” (Baptista 2002:249). This might be based on contexts like the pair in (139):

(139) *Dja-N graba di bu. / Dja-N graba di bo.
‘[you did not come to my birthday party] I am mad at you.’

I propose that the contrast in this pair depends on reasons different from the strictly syntactic arguments mentioned by the author.

These reasons are partly phonologically based: the preposition di ‘of’ is also a light word (recall that the bu form is Xº); hence, the phonological word that the two forms would build together – di-bu or di-u – can not bear the main stress of the sentence, which in the language falls on the right edge. Therefore, there are reasons to disallow this occurrence in the sentence-final position. We confirm this with the examples in (140), which contrast with the grammaticality in (138c):

(140) a. *N ten un livru pa-bu / pa-u.
b. N ten un livru pa bo.
‘I have a book for you.’

Resuming the question of whether subject clitics are heads in Agr or, rather, argumental clitics, a similar discussion has been central in different proposals on French clitics.

In De Cat 2005 some arguments are provided in order to show that the morphological analysis raises several unsolved problems, leading the author to defend the syntactic analysis. With some different details that are not relevant for the present
discussion, the same arguments can be presented for Capeverdean data – favoring, thus, the syntactic analysis of subject clitics in the language.

One previous note is necessary at this point: the most visible difference from French is that Capeverdean verbs do not display any overt agreement morphology; hence, redundancy (of person and number features) cannot be among the arguments against the morphological view of Capeverdean clitics.

Some of the arguments for French, though, apply to Capeverdean as well.

One of them is that subject clitics are available for syntactic operations (this would not be possible if they were affixes). Consider syntactic operations such as wh-movement (in subject questions no clitic stays in the subject position). This operation is illustrated in (141).

(141)  *Ki alunus ki (*esi) le kel-livru li?* 99

WH-pupils REL 3PL read DEM-book LOC

‘Which pupils read this book?’

For languages where subject clitics are grammatical agreement markers (agreement morphemes), rather than anaphoric agreement (incorporated pronouns, or syntactic clitics), it has been shown in Bresnan & Mchombo 1987 100 that they should be realized even when the argument (it is related with) is questioned (this is so in Chichewa’s subject markers, but not with object markers). This is not the case with subject markers in Capeverdean (141) – although a relative pronoun is there –, neither in French (142).

(142)  *Quels soldats; (*ils) sont partis?* (De Cat 2005)

which soldiers they have left

A different diagnostics for the grammatical agreement marker status, according to Bresnan & Mchombo 1987 and also De Cat 2005, has to do with locality: these

99 For the strategies and architecture of wh-interrogatives in Capeverdean, see Alexandre 2006, in preparation.
100 This study is cited in De Cat 2004; these authors consider data from Chichewa, and De Cat 2004 applies the same diagnostics to French, with the results favoring the anaphoric agreement hypothesis. I defend the same for Capeverdean.
morphemes cannot be associated with a DP which is in a higher clause, whereas what
the authors call anaphoric agreement needs not obey such a condition.\textsuperscript{101} In
Capeverdean, an associated lower clitic is possible in these long distance contexts.

(143) \textit{Djon, dja ten tres anu ki e, faze si kasa}.\textsuperscript{102}

‘Djon, there are three years now that he has built his house.’

One more argument against viewing French subject clitics as agreement morphemes is
that, if this view were right, there could not be any other preverbal material, that is, any
other clitics between the clitic and the verb.

This does not apply to French, where such clitics as \textit{en}, \textit{y}, object clitics and the
negation \textit{ne} may appear between the subject clitic and the verb. As in (144):

(144) \textit{Je ne t’ en veux pas.} \textsuperscript{(in De Cat 2005)}

1SG ne to-you of-it want not

‘I don’t begrudge you.’

This does not apply to Capeverdean either, where the negation \textit{ka} (and also the markers
\textit{ta} and \textit{sata}) as extensively shown earlier, appears between the clitic and the verb\textsuperscript{103}.

(145) \textit{E ka ta gosta di filmi.}

3SG NEG TMA like of movie

‘He won’t like the movie.’

\textsuperscript{101} At this point, and considering the proposal in Bobaljik 2006 on agreement being a morphological
rather than syntactic phenomenon (which has consequences on the traditional view of agreement as a
trigger for movement), some questions could be raised on the necessity of considering the
syntactic/argumental clitics as anaphoric agreement. I will leave this question open, since, in the present
work, the expression “anaphoric agreement” is not relevant.

\textsuperscript{102} An empty category in the embedded clause would be a variable, bound by a topicalized DP (variables
cannot be bound a c-commanding DP in an A position). This is not allowed in Capeverdean; the licit
variables in the language are bound by operators, such as \textit{wh-} expressions and quantifier DPs.

\textsuperscript{103} Also the status of \textit{ka} as a syntactic entity could be argued against, but as we may observe when
embedded non-finite contexts are at issue, the distribution of \textit{ka} is subject to syntactic constraints – an
embedded \textit{ka} only takes negation scope over the embedded verb, following the prediction that negation
cannot take scope higher than the clause containing it.
Finally, the most common argument in favor of subject clitics being grammatical agreement (morphemes) – and not, as I am defending for Capeverdean (anaphoric agreement) argumental clitics – depends on the information structure of clitic doubling contexts.

Under the morphological view, the topic interpretation (left-dislocation) of the NP is not expected – clitics are heads in Agr and the NP subject can surface on the canonical subject position, Spec,AgrSP.

On the other hand, under the view that subject clitics are argumental (syntactic entities), in the contexts where an NP and a clitic coexist the NP (co-indexed with the clitic) must be, according to the literature, left-dislocated (in a topic position – this means it surfaces in adjunction to a maximal projection: TP or CP).

There are various diagnostics to determine whether a given NP is left-dislocated (interpreted as the topic – in the sense of “what the sentence is about”) or not. De Cat 2005 contends that the decisive one is “not the presence of a pause between the XP and the rest of the sentence, but a combination of factors of which the most important are the presence of a stress (increased intensity) on the last syllable of the dislocated element, and pitch (melodic) prominence on that syllable”.

In Capeverdean, there is some difficulty in testing these phonological criteria to detect whether a given XP is topicalized or not, especially with an emphatic pronominal form. With a complex NP, however, at least the distinction of stress on its last syllable may be perceived. Check the contrast in (146).

(146)  a. Ami-\_N_ta bai pasa Natal ku nha mai.
      1SG.EMPH-1SG.CLIT TMA go spend Christmas with my mother
      b. Nha primu di Lisbon ta ben (p’-e pasa Natal ku nos).
      my cousin of Lisbon 3SG TMA come (for-3SG spend Christmas with us)

For (146a), where there is no clear pause, we see that the stress diagnostics must also confront some difficulty in determining the topic status of the XP pronominal. Hence, based on this diagnostics it is difficult to decide whether the subject clitic is a form of grammatical agreement (it would hold if Ami is not a topic and is realized on Spec,AgrSP) or a syntactic clitic (it would hold if Ami is left-dislocated).
In (146b), though, the stress distinction becomes clearer, which favors the left-dislocation proposal, allowing for the syntactic proposal for clitics.

Now consider (147).

   1PL TMA go cinema 2SG.EMPHATIC 2SG.CLITIC TMA go sea
   ‘We are going to the movies. (As for you,) you’re going to the beach.’

b. Abo bu ta bai.
   2SG.EMPHATIC-2SG CLITIC TMA go
   ‘You’re leaving.’

In (147a) we have the type of contrast traditionally presented in order to defend that with the pause there is topicalization, and without the pause there is not (the DP subject is in Spec,TP). As for (147b), it is clearer that the XP pronominal is left-dislocated; there is no context where the 2sg clitic cliticizes on *abo-u (which, given all we know about the phonological cliticization of subjects, could be expected).

Some further research is needed in order to determine whether the DP subjects in clitic doubling contexts are topicalized or not: this doubt is supported specifically by the Ami-N cases, just as in (146a).

Nevertheless, it seems reasonable for various independent reasons to propose that Capeverdian subject clitics are syntactic entities:

i) they are available for syntactic operations (such as wh-movement; and, we shall see further, for DP movements as well);

ii) they can be associated with a DP in a higher clause (they are not sensitive to locality);

iii) other clitics (negation ka and TMAs ta and sata) may appear between the subject clitic and the verb.

Still on the paused contexts, traditionally considered as the valuable diagnostics for a topicalized subject, the pause also marks the only possible environments (not common among my informants) where a full DP and a free form co-occur:
This reveals (as stated in Baptista 2002:257) that full DPs and free forms are in complementary distribution. Since in non-marked contexts any of them occupies the subject position (in Baptista 2002, the canonical subject position is Spec,AgrSP), they can only co-occur when one of them, the DP, is left-disclocated, leaving the subject position for the free pronominal form.

As for (148c), if one assumes, as I am proposing, that the clitic surfaces on T, the DP and the clitic (even with this being a syntactic entity) do not compete for the same position. Thus, if they are in complementary distribution, this is due to different reasons (for instance, Case reasons). As said before, however, it is not from the pause diagnostics that we may be sure Maria is left-dislocated, since there may be no clear pause.

Potential problems for the assumption that subject clitics surface on T are the following:

1. what happens to the subject theta-role (semantics)?
2. what happens to Nominative Case (syntax)?
3. what happens to the specifier of T?

One possible way out of this is to propose that, as argumental clitics, subject clitics are D's that project a DP; this would work in the fashion of what has been proposed in Galves & Abaurre 1996 for Brazilian Portuguese argumental clitics – although these authors propose this analysis to account for object clitics.

Assuming there is a trigger for movement (the nature of the possible triggers for movement will be discussed in 3.1104), we may consider that subject clitics are assigned a theta-role in their base-position (Spec,VP, in the case of transitive verbs), and move to

---

104 Given all empirical evidence, we must surrender to the fact that some elements are displaced from their base-positions. Moreover, minimalist concerns favor the view that this movement must be motivated; otherwise it is prohibited. One possible line of reasoning on this is the one proposing that syntax, by way of some abstract motivation, generates multiple outputs which are then filtered at the interface. The position of Capeverdean subject clitics – recall that there is no person/number agreement in the language – might well be accounted for in a similar fashion. I shall develop this in the next chapter.
T, where they absorb case (which, following a similar prediction for Agr in Holmberg 2005:537, is allowed, since: i) they are referential and ii) they head a chain.105

As an additional note concerning topicalized subjects, a peripheral question is related to individual-level predicate constructions. It has been assumed in the literature that these predicates can never appear in thetic contexts (where the whole sentence is focalized, with an information structure typical from answers to ‘What happened?’). More specifically, it has been assumed that subjects of individual-level predicates are always interpreted as topics. Therefore, the sentence in (149a) should be ill-formed, being the grammatical version (149b). As seen in the notations, though, this is not the case.

(149)  
a. Djon e riku.
Djon be rich.
b. * Djono, e i e riku.

I propose that this type of Capeverdean contexts – the sequence 3sg e + present copula e – displays phonological properties that may lead to wrong predictions. In (149a), for the left-disclocation hypothesis to be maintained, there should have been a phonological coalescence that “hides” the clitic (we cannot say that the “hidden” element is the copula, since what we hear is an /ε/ (the copula) and not an /e/ (the 3sg clitic). Moreover, we have seen before that copula deletion is allowed only in negative contexts, except for passive constructions.

105 A different possible problem for this positioning of clitics has been raised by Tjerk Hagemeijer (p.c.2006): how is the word order in negative imperatives derived? This might seem a problem indeed, since in these constructions we have the order NEG-clitic: Ka-u / Ka bu striba ku mi! ’Don’t you insist with me!’ (‘Don’t you push me!’”). This only holds, for obvious reasons, to second person clitics.

Baptista 2002:260 uses a negative imperative with the NP João, which is ungrammatical – * Ka João bai (’NEG João leave’). This intends to show that clitics and NPs have a different status with respect to negative imperatives. Although I find it difficult to figure out a context where a third person would be grammatical in a Capeverdean imperative – either with a clitic or an NP –, I suppose the line of reasoning on negation and imperatives in the present work favors, rather than the status of subject clitics as heads in Agr, the view that I am defending. If my proposal is on the right track, we do not have to resort to the raising of ka to Agr in imperative contexts, as the author does, since the possible reversed order in these constructions (declaratives: clitic - NEG / imperatives: NEG - clitic) obtains more naturally if:
a) subject clitics are syntactic entities, surfacing on a head adjoined to T;
b) only the nodes that are needed are projected.

We must assume the view that in imperatives the negation is of a different nature (which is not problematic under a semantic perspective): it is sentential negation, taking scope over the whole sentence. It is an adjunct and behaves like an adverbial. Since no temporal morphemes are allowed in these constructions, T is not projected. This accounts for the fact that the clitic does not need to raise to satisfy any case feature on T, and may stay in Spec,VP.
The phonological coalescence should, thus, be obligatory, since (149b) is prohibited.

In Pratas 2002, however (and as mentioned earlier), I have shown that the X° forms of the pronominals are disallowed with the copula \( e/era \). Hence, at least a free form is obligatory; this fact led me to consider that free forms behave in some cases as X°, having an ambiguous status (in these individual-level constructions they are not “free”), but this particularity is not relevant here. In this case, it seems to me that this incompatibility between the clitic and the copula is the true reason why (149b) is ill-formed. This incompatibility is proved in other person/number combinations (150a) and in a true left-dislocation construction with a free form (150b).

(150)  
a. * Ne spertu. / Mi e spertu.  
‘I am clever.’
b. N ka pode kumpra kusas ki Djon sata kumpra pamodi Djon, el e riku.  
‘I cannot buy the things that Djon is buying, for Djon, he is rich [and I’m not].’

I hope that the previous arguments, adapted from the analysis of French subject clitics in De Cat 2005, have clarified the syntactic, argumental status of Capeverdean clitics.

These arguments are as follows (as in page 138):

i) subject clitics are available for syntactic operations;
ii) there are no locality constraints between the clitic and the associated DP (these constraints are expected if clitics are agreement morphemes);
iii) some materials between the clitic and the verb are allowed (negation \( ka \) and temporal morphemes).

As stated earlier, I have not reached a conclusion on whether, even following the clearer diagnostics based on stress (on the last syllable of the dislocated element) rather than the one based on the presence/absence of a pause, the XP subject that may co-occur with the clitic is always topicalized.

One advantage of considering clitics as syntactic entities is that, besides accounting for previous problems concerning word order, this proposal accounts for the order \( dja+\)subject clitics. We shall return to this point shortly.
Since:

i) Capeverdean subject clitics are argumental entities (they have a syntactic status) that surface on a functional head;

ii) (as I am defending) in the IP domain there are no functional heads but T;

…the proposal is that:

iii) subject clitics have been copied in Spec,VP and merged on T (internal merge); from both outputs produced, rules and constraints filter the results at the PF interface, and the highest copy is pronounced.

It could be argued that this is a circular reasoning – such as: subject clitics are no reason for positing an Agr head; since there is no Agr, they must surface on T.

But recall that the absence of reasons for positing an Agr head largely transcend the clitics discussion (unless we do the circular reasoning in the opposite direction, positing an Agr head only because we assume that clitics are morphologic entities).

To put it clearer, positing an Agr head could be motivated either by:

i) the morphological and/or positional arguments discarded in 2.1;

ii) by the arguments involved in considering these clitics as agreement morphemes; the later arguments are also discarded by now.106

One more point to notice is the following: the proposal that subject clitics surface on a T head could lead to a conclusion in the line of the one proposed in Baptista 2002 – when a subject clitic occurs on its own (a context really productive in Capeverdean) there must be a referential pro occupying Spec,TP (in the author’s terms, Spec,AgrSP).

This should be so, for there is no subject DP to fill the canonical subject position. The main line guiding the present work, however, is that whenever there is no lexical or functional material to insert in a given functional position, this position is not projected. Hence, in these contexts (where the subject is a clitic) there is no Spec,TP.

In this subsection I hope to have shown, through empirical evidence observed in the light of theoretical concerns of economy, that there are no reasons to posit, for

106 More radical arguments against an Agr node in Capeverdean will be presented over the next chapters.
Capeverdean, a split of the maximal projection, IP, into an Agreement node and a Tense node.

Finally, in the next subsection I will develop my previous proposal on Capeverdean functional structure.

2.4.2 Capeverdean TP: an almighty projection?

A corollary of the evidence presented up to this point is that TP is the functional projection in Capeverdean inflectional domain. There are tense and aspect marking morphemes (as shown in the previous sections, mood does not seem relevant in the language; either way, TMA is a label that can apply to these morphemes, given the still obscure feature complexities involved) to be analyzed as instances of recursive head adjunction, which straightforwardly accounts for the string-adjacency facts.

This proposal respects the core assumptions in minimalism, which could, for the present topic, be subsumed in the terms of the Real Minimalist Principle (Thráinsson 1996:261), which guides language acquisition and parameter setting: “Assume only those functional categories that you have evidence for.”

For Capeverdean, more functional categories do not add clarity. In order to account for the semantic facts discussed in subsection 2.3.1, and also for the difficult task of mapping these onto a hierarchic functional structure, I propose that, in the language:

i) Subject clitics and Temporal features have been merged with Tense;

ii) The negation marker ka and the copula forms e (present) and era (past) also surface under the label T, as it has been shown by their behavior with respect to other functional morphemes. Namely:

   a) ka shows the typical properties of a head and it occurs in a position previous to TMA morphemes, except for djia;

   b) it occurs pre-copula era, but shifts positions with the present copula e; e, which encodes no content other than Pres, is preferably null in the presence of ka;

   c) the different behavior of e vis-à-vis its past counterpart, era (and also foi), is due to the fact that it is morphologically lighter.
Given all this, I propose for Capeverdean the TP structure in (151) for a sentence like
*Djon ka ta kume katxupa* ‘Djon does not eat katxupa’:

(151)    
  TP
  |    
  Spec, TP  T’
  |    
  Djon_i
  |    
  T  VP
  |    
  ka
  |    
  T  Spec, VP  V’
  |    
  ti
  |    
  V
  |    
  DP
  |    
  kume
  |    
  katxupa

In (152) (next page) we have the representations for copula constructions, where the
forms *e* and *era* are nothing else than temporal markers. The complement for *T* is not a
VP but rather a lexical projection headed by some predicate other than a verb (just like
in small clauses). It can be an AP (*spertu* ‘clever’), a DP (like in *Djon e mediku* ‘Juau is
a doctor’) or a PP (like in *Kel-li e pa bo* ‘this is for you’).

In (152a) we have the representation for a sentence like *Djon ka era spertu*
‘Djon was not clever’, where the temporal specification (Past) must be overt.

In (152b) and (152c) we observe the morphological transformation operated.
From the ill-formed *Djon ka e spertu*, we obtain *Djon (??e) ka spertu* ‘Djon is not
clever’: in (152b) we observe the trading of positions between *ka* and *e*, which is
possible only with this Pres form, as said earlier, due to its morphological lightness (as
opposed to *era*); then, with *e* in this position, being its temporal (Pres) content merged
with the one spelled by *ka*, its deletion is allowed (152c). Recall that this is attested by
the impossible *e* deletion in contexts with no *ka*.

Confront this with the absence of the copula, even in non-negative contexts, in
passives: *Djon tradu di trabadju* (Djon taken of work) ‘Djon has been fired’. This adds
one more argument for the present analysis: the necessary temporal specifications are
given in passive morphemes (*-du* (Pres) and *-da* (Past)), therefore the copula is not
needed at all (hence, it is not allowed) in these contexts.
In these diagrams I am abstracting away from the possible topic nature of a DP subject in individual-level constructions.

If it were the case here, it would mean that in *Djon ka spertu* we would have the deletion both of the copula and the 3sg clitic *e* – that is, for the representation before any trading of positions and any imaginable phonological coalescence (in the case that the DP was topicalized) the diagram should include both *e* forms, and also *Djon* in a peripheral position. Something like: *Djon, e ka e spertu*. Nevertheless, it is clear that the coalescence should occur after the trading of positions between the negation and the copula. The diagrams above do not have the intention of being decisive with respect to this possibility.

Testing with other environments, where this phonological similarity between the copula and the clitic would be excluded, we get to no further conclusion. Check the following: with a 2sg clitic (ia), in an evident topicalization context, the copula is preferably null; in an extraposed construction (ib), where there is no matrix subject, the copula is obligatorily overt and in a pre-negation position.

(i) a. *Abo, bu (??e) ka spertu.*
   ‘You, you are not clever.’

   b. *(E) ka faxi pa-m dexa di fuma.*
   ‘It is not easy for me to quit smoking.’
Now that I have described how the negation *ka* and the copula *era* can be likewise accommodated by T, there are still two particular phenomena that might be argued to be problems for this proposal (on TP being the only functional projection): the distribution of the TMAs *-ba* and *dja*.

The first could be a problem since it appears to the right of the verb, which could be taken as evidence for verb movement – and, thus, for the positing of distinct functional heads. The latter could be a problem since it surfaces in a pre-subject clitic position, which could also seem a motivation for projecting distinct functional nodes. I shall address each of these potential problems in the next two points.

### 2.4.2.1 Capeverdean T and the *-ba* problem

The postverbal affix *-ba* could, as a matter of fact, be considered one potential problem for the claim that there is no evidence for V-to-I movement in Capeverdean (Pratas 2002). In other words, it seems positive evidence for V-to-I. Let us consider one of its occurrences.

(153)  

\[ N \, t a \, f u m a b a \, t x e u. \]

‘I used to smoke a lot.’

The postverbal position of *-ba* is analyzed in Baptista 2002, not as a morphophonological consequence of there being V-to-I movement in this language, but as the very trigger for the syntactic movement of the verb to T° (Baptista 2002:203).

However, there seems to be no other empirical evidence to consider V-to-I movement in Capeverdean (Pratas 2002) – this issue will be addressed in greater detail in Chapter Three.

In Costa & Pratas 2003 it has been shown that the distribution of *-ba* is no reason either. Following Halle & Marantz 1993 and Embick & Noyer 2001, we have proposed then a different interpretation for this fact (without resorting to V-to-I movement): postverbal *-ba* involves lowering of *-ba* to V.
Consider the description of Lowering: “A zero-level element trades its ‘head-to-complement’ relation with its complement for a relation of affixation to the structural head of its complement.”

This alternative view on postverbal -\textit{ba} makes interesting predictions.

Since Lowering of T to V is, thus, an operation that occurs after syntax, in the morphological component of the grammar:

\begin{itemize}
  \item[i)] no V-movement is needed;
  \item[ii)] adjacency between -\textit{ba} and the root is required. It is therefore predicted that there is no material in between T and V, just like in English with the morpheme -\textit{ed} (Bobaljik 1995).
\end{itemize}

As we can see in the representation in (154), for a sentence like \textit{Djon sata kumeba bolu} ‘Djon was eating a cake’ (PastPro) the adjacency requirement is satisfied, since the subject DP, generated in Spec,VP position, is already in Spec,TP when this Lowering operation occurs (the boldface font marks the moved constituent / morpheme).

\begin{figure}
\centering
\includegraphics[scale=0.5]{representation.png}
\caption{(154) TP}
\end{figure}

The subject, \textit{Djon}, moves in syntax (internal merge) for Case reasons, and the morpheme -\textit{ba} lowers after that, in the morphological component.

In the view of this, one could ask some predictable questions:

\begin{itemize}
  \item[i)] why is there a specificity for -\textit{ba}, and also the passive markers -\textit{du} and -\textit{da}?
  \item[ii)] in particular, why are only these subject to Lowering, and not \textit{sata} or \textit{ta}?
\end{itemize}
Alec Marantz (p.c.) and Michel DeGraff (p.c.) point out that the necessity of lowering is indeed morpheme-specific. In English 3sg -s and past -ed lower, but future (will) does not. This could mean that Capeverdean would be one more instance of a mixed language (for mixed systems, with free and bound morphemes, see Cinque 1999).

We already know, from 2.3.2, that -ba exhibits some specific properties: for instance, it is the only TMA morpheme that may co-occur with ta and sata, as we can see in the template in (39), here repeated as (155):

(155) a) V (or V + ø )  
   b) ta V  
   c) V-ba  
   d) ta V-ba  
   e) 1) sata+V (inland of Santiago); 2) sta+ta V (other Sotavento areas)  
   f) 1) sata + V-ba; 2) stab + ta V(-ba)  
   g) ta stab + ta V-ba  
   h) * ta sata + V  
   i) * sata + ta V

When -ba co-occurs with ta or with sata, it is empirically attested that it is -ba that appears to the right of the verb; there is, thus, no question about being -ba the one in a lower position, relatively to the others. The question could arise when -ba is not present.

Before proceeding with the description of all the details involved, let me represent my proposed order for the different morphemes in (156). It can be argued that the incompatible ta and sata may occupy the same slot in this hypothetical tree, since they are in complementary distribution. The same happens with the three morphemes that are subject to Lowering, which are, obviously, in complementary distribution as well: -ba, -du and -da. This complementary distribution, however, is not a consequence of the adjunction configuration, but rather of the semantic incompatibilities between the mentioned morphemes: the possible temporal contributions of ta are incompatible with the ones possibly brought by sata; the temporal contributions of -ba, -du and -da also display obvious incompatibilities.
Recall that *dja*, the negation marker *ka* and the copulas *e* and *era* have specific properties of their own that show they occupy a higher\(^{109}\) position than the one reserved to *ta* and *sata*; *dja*, as a functional morpheme (its status as an adverb is not at stake here), is the only one that may appear to the left of the subject clitic and also to the left of negation *ka*, which shows that it must be the highest of them all; the negation *ka* must occur at the same level as the copulas *e* / *era*, for the present form *e* (for reasons of morphological lightness) can trade positions with *ka* and, then, be deleted, as we have seen in (152); this level must be one of the highest – after *dja* and the subject clitic (in this order) –, since in contexts with *ta* and *sata, ka* appears to the left of these.

As for the specific *-ba*, and also the passive morphemes *-du* and *-da*, it would still be difficult to explain with recourse to any V-to-T movement: why should V move upwards to the lowest T position, getting *-ba* as its affix, and not to the higher *ta* or

---

\(^{108}\) The discussion around this position for clitics follows, in the next point. Nevertheless, I include them in this representation in order to get the whole picture of the functional domain.

\(^{109}\) I am considering the relative position of functional morphemes as higher and lower, in spite of them being in successive head adjunction. For the present purposes, this is a matter of naming their relative – linear – position in the string adjacency that characterizes their distribution. In other words, “higher” means to the left in the sentence and “lower” means to the right.
sata? Notice that the affixation of these morphemes never occurs, even when -ba is not present.

In Chapter Four we shall see that there is more to the morpheme -ba, since its presence has consequences on the selection of the categorial status of the object pronominal. Before that, however, let me add some important elements to the functional domain.

2.4.2.2 Capeverdean T and the dja-clitic problem

In 2.3.2.4, the ambiguous status of dja has been discussed: it can be in preverbal position and have a temporal morpheme interpretation – contributing some temporal location and/or perspective –, and it can appear in sentence-final position and have an unambiguous adverbial status.

Furthermore, preverbal dja shows some other idiosyncrasies of its own.

Whenever the subject is a clitic form, this one appears to the right of dja, and it may cliticize phonologically, resulting in the form dja-clitic; whenever the subject is a DP or an emphatic or free pronominal form, the order is reversed: subject + dja. The relevant data are presented in (157).

(157) a. Dj’-e sabe. / *E dja sabe.
   DJA-3SG.CLITIC know / 3SG.CLITIC DJA know
   ‘He/she knows.’

b. Dj’-e sabe dja. / *E sabe dja.
   ‘He/she knows (it) already.’

c. Djon dja / * Dja Djon sabe.
   DP DJA know
   ‘Djon (?already) knows.’

The order constraints between the preverbal dja and the different subject categories (the emphatic form of the pronominal behaves like a full DP) could seem a problem for the main proposal in this chapter: all functional morphemes surface on head adjunction
under T. How come that subject clitics surface to the right of the TMA marker \textit{dja} and to the left of other TMAs, such as \textit{ta} or \textit{sata}\footnote{Recall that, as said earlier, \textit{dja} could be analyzed as an argumental word with an aspectual meaning and surface on Spec,TP. In English locative inversion (Bissel 2004), a locative PP can appear in Spec,TP. In Holmberg 2005 it is argued that certain referential adverbials (locative, temporal, instrumental – but not manner or reason) can check the EPP. The first and main reason for my rejecting of \textit{dja} on Spec,TP is that this positioning of a temporal marker has only one purpose: satisfy the EPP. In the next chapters it will become clear why the EPP may be a point of parametric variation and, as such, have a negative value in Capeverdean (any syntactic operation that serves only to satisfy the EPP is, then, unnecessary; thus, it is not allowed).}

Crucially, what I have called the \textit{dja}-clitic problem is not a problem for my proposal; in fact, it brings some more arguments that sustain it. The reasoning goes as follows:

A. We know that…

i) (for some reason that I have not yet been able to disclose) \textit{dja} always appears before all other functional morphemes that may co-occur with it (\textit{ka}, \textit{sata}\footnote{Recall that preverbal \textit{dja} and \textit{ta} never co-occur (\textit{ta} and \textit{sata} never co-occur either); I propose that this constraint is grounded on semantic reasons: their respective contributions are not compatible. Recall also that the other morphemes have order requirements of their own: \textit{ka}; \textit{ta} / \textit{sata}; -\textit{ba}.}) and also before the subject clitics; a possible reason could be related to scope: \textit{dja} (when not an adverb) is a temporal marker that selects the whole sentence as its scope domain\footnote{This would account for the fact that when there is a subject DP also realized – that is, the order DP + \textit{dja} + clitic – the reading for \textit{dja} is more like a temporal morpheme (with all the previously discussed consequences, given the aspectual properties of the predicate in question, etc.); when the subject is a DP, with no clitic, the reading of \textit{dja} is more like an adverbial. We can see this contrast in (157): \textit{dja} + subject clitic has no adverbial reading (157a); \textit{dja} + subject clitic + sentence-final \textit{dja} has an adverbial reading given by the last \textit{dja} (157b); subject DP + \textit{dja} may have a light adverbial reading (157c).}

ii) in certain phonological conditions, clitics may cliticize to their own left, onto some different hosts (preposition \textit{pa}, complementizer \textit{ma}, among others).

B. If we assume that…

i) Capeverdean subject clitics surface in head adjunction under the functional label T as well (which does not result in their losing their argumental status)\footnote{Although, in this case, there is no temporal information involved, what matters is their syntactic behavior, according to which this must be in fact their surface position (since they are base-generated in Spec,VP, where they are assigned a subject theta-role, moving upwards to a head position).};
ii) when *dja* is realized, the position for clitics immediately follows *dja*, for, say, some reason related to its temporal reading having to have scope over the whole sentence.

C. Then it is expected that …

i) whenever the subject, instead of $D^\circ$, is a DP or an emphatic form (which also behaves as an XP), *dja* surfaces to the right of it;

ii) in these contexts where a DP is the subject, *dja* may have a light adverbial reading if no subject clitic is realized; if a subject clitic (co-indexed with the subject DP) appears, the *dja* reading is clearly more like a temporal morpheme (as said before, I will leave for a further discussion the question of whether this subject DP is left-dislocated/topicalized).

D. An extra argument:

The fact that there are no maximal projections between the subject clitic position and VP, and the absence of such morphemes as *ka*, *ta/sata*, allows for a strict adjacency between the clitic and the verb. This may be the reason why, with verbs like *bai* ‘go’ or *papia* ‘speak’ (and when there are no TMAs like *ta* or *sata* in the sentence), the subject clitic $N$ turns into $M$, for it is in the necessary conditions to absorb the labial feature from the adjacent consonant.

The final representation in this chapter is the one including the relative positions of *dja* and the subject clitic. This diagram, in (159), stands for the sentence in (158).

(158)  
*Dja-N sabe ma bu ta ben.*

‘I know that you will come.’

(159)  
```
            TP
              |
            T   VP
        |
    T   V   CP
    |     |
    Dja  sabe
        |
        N ma bu ta ben
```
In this last descriptive point of the present chapter I have argued that the temporal functional morpheme *dja* (whose distribution has been largely illustrated and described in the subsection 2.3.2.4) and the subject clitic both occur in head adjunction under the label T.

Their strict adjacency allows for the clitic leaning on *dja*, a phonological host to the left. Recall that this positioning of the subject clitic (Dº, which has absorbed its theta-role in its base position) on a functional head does not change its argumental status.

Whenever there is a subject DP, it surfaces on a pre-*dja* position. In the contexts where there is no clitic, the DP is on Spec,TP and *dja* may have a light adverbial reading. When there is also a clitic, *dja* appears consistently post-DP and pre-clitic, and may seem like a temporal/aspectual morpheme. The question of whether in these latter contexts the DP is necessarily left-dislocated must be addressed independently (it is related to the syntactic status of the clitic).

### 2.5 Conclusion

The argumentation provided in the present, extensive chapter makes it hard to sustain, for Capeverdean, an immense array of functional categories and the surfacing of each morpheme that marks temporal reference under a distinct label, such as Agr, Asp, Tense, Mood, etc. I hope to have shown through Capeverdean relevant examples that:

i) there is no evidence – provided neither by any overt agreement morphology or by any sort of positional arguments – for positing an AgrSP node;

ii) contrary to the universal functional hierarchy analysis (as in Cinque 1999, which predicts that adverbs behave uniformly cross-linguistically, surfacing on specifiers of given functional heads), and following a more flexible approach (Ernst 2002, Costa 2004), Capeverdean adverbs are adjuncts; they appear in sentence-initial and sentence-final positions (except for biclausal structures, where certain temporal adverbs may appear between the two TPs).
iii) TMAs appear in sequences unbreakable by any intervening element (no XP may disrupt their adjacency). The same morpheme may instantiate different meanings (I hope to have shown this extensively, crossing theoretical concepts and relevant data), depending on distinct factors, namely: a) combination with other markers, expressions and discourse information, and b) semantic properties of predicates. As it has been proposed in Pratas 2004 and also in Costa & Pratas 2005, TMAs surface in successive head adjunction under the label T – this includes the negative marker *ka*, which displays a head behavior, and also the subject clitic. There is no evidence for positing additional functional heads. It is therefore expected that their interpretation, as it happens in other cases of adjunction, depends on their lexical properties and the composition of the other adjuncts.

The fact that subject clitics also appear on this functional position brings one more possible environment leading to the argument of Spec,TP being only projected when needed. This proposal will be discussed in greater detail especially in Chapters Five and Six.

In the next chapter the Verb domain will be addressed. Besides describing the relevant properties of its internal structure, the main claim will be that there is no empirical reason to consider that there is V-to-I (in the present case, V-to-T) movement in Capeverdean.
Chapter three

Morphology versus abstract features

In Chapter Two, the Capeverdean inflectional domain has been addressed. Grounded on empirical facts, the main proposal has been for the pieces of temporal inflection to surface on multiple adjoined T heads. Besides accommodating all the TMA morphemes, Capeverdean TP is sufficient to account for the negation head *ka*, the copulas *e* (Present) and *era* (Past) and, finally, subject clitics.

The next step is to go down in the structure and analyze some proposed abstract relations between verbs, arguments and Tense. The ultimate goal in this chapter is to show, in a different level and through other problems and operations, that the PF interface plays a crucial role in the language. Considering the lack of overt verbal agreement morphology, and proceeding through the complex implications of Case and agreement (phi-) features, and also of T- and V-features, the main claims will be:

i) there is no empirical evidence to propose a V-to-T movement;

ii) the positions in which subjects are spelled out are determined, not by strictly syntactic (abstract) features, but rather by rules and constraints at the interface level:

a) for them to surface on Spec,TP, the PF interface is at stake (imposing strict adjacency conditions between the TMAs and the verb);

b) for the definite subjects of unaccusative verbs to be forbidden at their base position (hence, surfacing on Spec,TP as well), there must be some constraint at the LF interface level that rules out definite DPs in this position\(^{114}\).

\(^{114}\) We may assume, following Beletti 1988, that Partitive case is assigned in these contexts and that this type of inherent case is only compatible with indefinite NPs.

Also when other DPs are at stake in this post-verbal position, Partitive case seems to be related to some semantic features of this DP: consider the reasoning in De Hoop 1992, according to which Partitive is the weak case, producing non-specific readings of the object, whereas Accusative is the strong case, producing specific readings of the object.
Some of the arguments that sustain these views have been introduced in the previous chapter, namely: a) there is an alternative solution (a Lowering operation) to the affixation of \(-ba\), which, thus, is no diagnostics (or trigger, for that matter) for V-to-T movement (which is traditionally referred to as V-raising); b) the proposal that subject clitics are Dº that project DPs, absorbing their theta-role in their base position and then surfacing on a T head, which does not interfere with their argumental status (they are, for instance, available for syntactic operations).

In the present chapter, an extended introduction of what may happen around and within VP is included as a motivation for this claim. In Chapter Four other verb movements, such as movement to a higher V in a VP-shell configuration, will be considered.

The present chapter is organized as follows. In section 3.1 the relation between VP and the licensing of DPs will be at issue. I shall summarize some evolving approaches on Case, its relation to agreement and also to some minimalist theoretical tools regarding A-movement, such as feature checking and Agree (3.1.1). These will be presented under a critical view, particularly following some arguments in Bobaljik 2006.

The questions concerning the specific licensing of subject DPs will also require some notes on the EPP (3.1.2), although this will be discussed in greater detail in Chapter Five. Empirical evidence for instance from German (Wurmbrand 2001) and from European Portuguese (Costa 2004), among others, show that EPP is not universal; the more radical proposal in Epstein & Seely 2006 is that EPP does not exist.

My own proposal on Capeverdean DPs will be presented at last (3.1.3), positing that their distribution depends on a morphologically motivated filter, at the interface with phonology.

In section 3.2 I will address some verb raising issues: in subsection 3.2.1 I summarize some remarks on the motivations for verb raising; in 3.2.2 Capeverdean V-to-T is under discussion, as well as the classical diagnostics for this transformation: the behavior of floating quantifiers (3.2.2.1) and adverbs (3.2.2.2); finally, in subsection 3.2.3 I will recall and put into perspective some arguments around verb raising and the affix \(-ba\).

In the end of the chapter (3.3) a conclusion will be provided.
3.1 The licensing of DPs: a Case and agreement matter?

In this introductory, more theoretical section I shall address some topics on the licensing of DPs, arriving to a proposal in order to account for Capeverdean DPs, especially subjects. This proposal (in 3.1.3) is not strictly about abstract Case or agreement (whose implications are described in 3.1.1), in a feature-driven theory (which is developed in 3.1.2), but rather about the relation between syntax and PF (that is, the interface syntax-phonology).

Consider the Capeverdean sentences in (160).

(160) a. *Djon ka ta papia ingles.
   ‘Djon does not speak English.’

b. *Ka ta papia Djon ingles. / *Ka ta papia ingles Djon.

c. *Ka Djon ta papia ingles. / *Ka ta Djon papia ingles.

d. More sentu di algen na fomi di [19]46. / Sentu di algen more...
   ‘Hundreds of people have died in the 1946 famine.’

   ‘My grandmother died in 2002.’

f. Ten tres katxor na nha kasa.
   ‘There are three dogs in my house.’

g. *Tres katxor ten na nha kasa. / ???Tres katxor sta na nha kasa.
   ‘There are three dogs in my house / three dog BE in my house

h. Sata txobe txeu.
   ‘It is raining a lot.’

The position of Djon, the external argument of a transitive entry of the verb papia ‘speak’, is problematic in (160b) and (160c): in the two sentences in (160b) Djon is in a postverbal position (pre-object and post-object, respectively); as for the sentences in
(160c), in the first Djon is post-negation and pre-TMA (it is within the functional domain) and in the second it is supposedly in Spec,VP (its base position).

With the unaccusative verb *more* ‘die’, the postverbal position for the internal subject *sentu di algen* ‘hundreds of people’ is okay, as in (160d), whereas this same position for the subject *nha dona* ‘my grandmother’ with the same verb is disallowed in (160e) – notice that there is an indefiniteness/definiteness distinction between the two.

As for the existential construction in (160f), the postverbal position for the subject *tres katxor* ‘three dogs’ is the only one available – confront with (160g), where the counterpart with *sta* ‘be’ is also problematic. Finally, (160h) shows a weather construction, where everything is fine.

No overt expletive is in the subject positions in (160d), (160f) and (160h).

We will return to these sentences in subsection 3.1.3. Before that, however, I will summarize some of the theoretical tools trying to account for subject positions as motivated by Case and / or agreement licensing.

One of the purposes of this short review is to align some core questions on the true role of Nominative Case:

i) is it the motivation for syntactic operations, such as movement?
ii) how is it related to agreement?
iii) how is it related to the grammatical function (a primitive in the grammar) that we call subject?

As we shall see, this is a matter of great dispute, with recent studies casting more doubts on these dependencies and interactions.

To consider, for instance, that Nominative feature checking is a trigger for movement of the DP has been, at a certain moment in generative studies, a great resort to explain the contrast between sentences like: *Djon ka ta papia ingles / *Ka ta Djon papia ingles. The argument would be that, in the latter, the DP in its Spec,VP position has not checked its Nominative features; in the former, where the DP has moved to Spec,TP, this problem is solved.

Nevertheless, we shall see that, contrarily to some traditional analyses on Nominative Case effects, in some languages and under certain conditions, a DP does not necessarily have to be displaced from its base position in order to be a legitimate
syntactic object. Hence, we must concentrate on some other language-specific properties and constraints in order to determine whether this is the relevant violation in the ill-formed sentence at stake.

Moreover, one must decide whether the interpretation of Case as an abstract entity (sometimes independent from any morphological markings for case) is of some utility, let alone the problematic evidence about its exact status.

For instance, the study in Zaenen, Maling & Thráinsson 1985, empirically based on Icelandic data, defended that the relevant licensing theory be formulated in terms of grammatical functions (subject, object, etc.), rather than in terms of Case theory.

I shall propose for Capeverdean that the licensing of subject DPs in certain word order configurations is conditioned by rules and constraints at the interface levels. The necessary movement through which they appear in certain positions must, however, be motivated (otherwise, as was the case in the GB era, unconstrained movements would give way to an overgeneration of products). Given the point where theoretical studies are at the present moment (in other words, despite all the doubts and inconclusive proposals on different languages and environments), there seems to be no unobjectionable trigger for movement of DP subjects in Capeverdean other than Case.

After presenting some of the steps primarily involved in the permanent discussion around Case assignment (3.1.1), in 3.1.2 I will present some evidence against the obligatory movement of subject DPs (grounded on examples and arguments provided in Wurmbrand 2001 and Costa 2004) and against the universality of the EPP; in 3.1.3 I will present my own proposal, grounded on Capeverdean data and some already mentioned grammatical facts.

### 3.1.1 Agreement and Case studies

According to the Case Filter, every overt DP must be assigned Case\textsuperscript{115}. The questions around the way in which this process applies, which particular category in a given

\textsuperscript{115}This idea – that at the time helped explain the distribution of referential null subjects of infinitival clauses (PRO) – has been presented in Chomsky 1980, 1981 (who attributes it to Jean-Roger Vergnaud).
configuration is responsible for it and how it is related to morphological case have, however, been subject to a broad debate.

In the two first points in this subsection I will present a summary of the steps involved in considering Case assignment a necessary condition for the licensing of DPs. Firstly it was assigned under government (in the GB era, this was a core relation in the grammar); over the years it became a matter of abstract features (in the sense that they do not have any morphological realization\textsuperscript{116}), needing a few more concepts and machinery (in the MP era, the core relation is Agree).

In 3.1.1.1 a general overview is provided. In 3.1.1.2 some questions around this transition with respect to Accusative Case are presented – although Accusative Case is not relevant for the present work (this short point is, say, merely informative). In point 3.1.1.3 a critical view on this mixing of abstract Case (syntax proper) and agreement (morphology, after syntax) is presented.

\textbf{3.1.1.1 From government to Agree}

In the standard Government and Binding (GB) era, following Chomsky 1981, Nominative Case was assigned under government.

In Chomsky 1986 (Barriers) the notion of m-command (crucial to the definition of government, a core relation for GB) is the following:

\begin{itemize}
  \item A m-commands B iff A does not dominate B and every maximal projection (XP) that dominates A also dominates B.
\end{itemize}

As for c-command, the relevant definition is given below.

\begin{itemize}
  \item A c-commands B iff A does not dominate B and the first branching node that dominates A also dominates B.\textsuperscript{117}
\end{itemize}

\textsuperscript{116} This notion of abstract Case allowed for, for instance in Icelandic, some (abstractly) Nominative DPs (the subjects of finite clauses) and some (abstractly) Accusative objects to be morphologically realized as any of the four available cases: nominative, accusative, dative (this is the so-called quirky case for subjects) or genitive.

\textsuperscript{117} This definition of c-command has also been referred to as strict c-command. In this context, the wider definition of c-command – A c-commands B iff A does not dominate B and every X that dominates A also dominates B – predicts both strict c-command (whenever X is the first branching node, like Z') and m-command (whenever X is every maximal projection, like ZP).
The relevant definition for government is as follows:

A governs B iff

i) A is a governor; and

ii) A m-commands B; and

iii) no barrier intervenes between A and B.

Maximal projections are barriers to government.\(^{118}\)

Governors are heads.\(^{119}\)

Thus, the DP subject, on Spec,TP / IP, would get its Nominative Case from Tº / Iº, in a proper specifier-head configuration. On the other hand, also Accusative Case was assigned under government, but in a head-complement configuration.

Later (in particular in Chomsky 1991), this asymmetry between the strategies for assigning Nominative and Accusative has been corrected, with the proposal for Accusative also to be assigned under a specifier-head configuration: the responsible for this would be the functional projection AgrOP. For languages like English, where there is no object shift and no V-to-T movement (the objects surface post-verbally), it has been assumed that this movement of the Accusative DP from its object position to Spec,AgrOP is covert (it occurs at LF, that is, after Spell-Out). Thus, Accusative Case would be assigned under spec-head agreement as well.

Under Minimalist Program (MP) proposals, though, grammatical relations are no longer defined through invariant structural relations.

One crucial example of some facts unexpected under the previous (GB) assumptions is the so-called Quirky Case. This generic name includes Dative Subject Constructions in various languages, namely Japanese and Icelandic, in which DPs that show subject properties (such as being able to bind subject-oriented reflexives or to control the missing subject of an adjunct clause; or such as surfacing at Spec,IP) show Dative Case, instead of Nominative.

\(^{118}\) IP will be excluded – that is, IP will not be considered a barrier for government.

\(^{119}\) Later, the notion of minimality is introduced (basically, if there is a potential governor for B, say Z, that c-commands B and does not c-command A, A cannot govern B – when there is competition between two potential governors, the closer one wins), and also maximal projections will be potential governors, as long as they are co-indexed with the governee; this is antecedent-government, whereas the one defined here is head-government. Either antecedent-government or theta-government (A theta-governs B iff A governs B and A theta-marks B) must take place for proper government to occur.
With the progression into MP era, Case has acquired the quality of a formal feature, which shows some idiosyncrasies of its own.

This abstract feature has been connected with agreement relations. Concerning Nominative, this agreement relation is established between a functional head T and a lexical head N. This confusion between Case and agreement (neither of them having a necessary morphological expression) has been one of the big puzzles: why does T (which is associated with verbal material) bear a feature inherently associated with nominals?

One simple solution has been to consider that Nominative on N is Case, and Nominative on T or V is agreement. A different one proposes agreement as being the necessary relation established between the features themselves; that is, between an uninterpretable instance of a given feature and an interpretable instance of the same feature. Thus, Nominative Case is a T feature that is interpretable on Tense, but uninterpretable on N (or D, in the case that we are considering a DP instead of NP).

These two instances of this T-feature must enter an agreement relation for the derivation to converge. The configuration necessary for this agreement relation to obtain is the next big puzzle.\[120\] In a parallel, φ-features (phi-features, for person and number) are interpretable on N/D but uninterpretable on T or V; the φ-features on T and V must agree with their counterparts on N/D.

Case on N/D is universally marked as [-interpretable] – uninterpretable –, and thus is required by Full Interpretation, an LF interface condition, to be properly licensed: checked, deleted and erased. To be deleted means that the checked feature is rendered invisible at LF but still accessible to computation. When a deleted feature becomes inaccessible to computation, it means it has been erased (Chomsky 1995:280).

A further idiosyncrasy of Case is the following: as a feature, it can be strong or weak.

Both the possibility of it being checked after Spell-Out (in the LF branch) and the obligatoriness of it being checked before Spell-Out (which renders the operation visible to the PF branch) follow straightforwardly from this critical property of strength:

\[120\] T-features are also uninterpretable on V, which must also enter an agreement relation with T. In some languages this triggers V-to-T movement, in other languages (like English and, I will argue, Capeverdean), this head-to-head relation is satisfied without movement. Notice that, according to some views, this Agree relation is also on the interest of T, whose T features are unvalued and get their value from the finite Verb.
- either the Case feature is weak (being legitimate at PF, the derivation converges);
- or the Case feature is strong (which makes it uninterpretable at PF and, unless it is checked and deleted before Spell-Out, causes the derivation to crash).

The computation has two available operations: Merge and Move. In other words, Move is a form of Merge – thus, in a sense Merge is the only real operation –, since it copies an element that is already part of the structure and merges it in a different position; this is called internal Merge, as opposed to external Merge (when some new element is added to the structure, which is always built upward).

According to the Theory of Movement, movement is triggered (Chomsky 1995, 2000, 2001; Alexiadou & Anagnostopoulou 1998; Pesetsky & Torrego 2001 and subsequent works). Trying to refine the whole process, we might say now that the first ingredient of movement is Agree, a relation between an uninterpretable feature ($uF$) of a probe category H (a functional head) and a corresponding feature of a goal category G (the potential checker) within some local domain (phase). These features must be active.

Phi-features ($\phi F$) make a semantic contribution when found on DP or CP, but make no semantic contribution on T. “Although it is difficult to know why lexical items bear uninterpretable features [$uF$], their existence is a fact.” (Pesetsky & Torrego 2001:359)

Following Chomsky 2000, under certain circumstances the Agree relation that is the trigger for movement may allow for a Case/agreement licensing without movement as well (notice that, under this view, Agree does not involve any copying of features). The abstract relation established between the probe and the goal is, then, sufficient. Except for what one understands by local, this relation might be considered similar to the head-government in the GB era.

In a nutshell, a description of the evolution of the notions of Case and agreement has been presented. The relevant idea is that the minimalist proposal of viewing Case as an abstract feature and the notion that agreement is a relation between uninterpretable and interpretable instances of the same feature has similar results as the old notion of government.
In the next point I will summarize some specific questions regarding these views on Accusative Case. This comes as an informative note on the side, mainly to give some environment details to the big picture of the above mentioned discussions.

3.1.1.2 Direct Objects and Accusative Case
Traditionally, the conditions for Accusative case assignment are the following:
- \( \alpha \) assigns Accusative case to \( \beta \) if
  i) \( \alpha \) is V or P (not N or A);
  ii) \( \beta \) is either the complement of \( \alpha \) or the specifier of the complement of \( \alpha \);
  iii) \( \alpha \) and \( \beta \) are adjacent.

For the present description, only abstract Accusative Case is relevant.

Before proceeding, it is important to notice that the subjects of unaccusative verbs are likewise assumed to be generated as a complement of V – as internal arguments. These verbs are out of the whole Accusative discussion, though, since, not being able to assign Accusative Case (this is expressed in Burzio’s Generalization\(^{122}\)), their internal argument has to move, say, to Spec,TP, where it checks its (Nominative) Case feature or, in the old terms, it satisfies the Case Filter.\(^{123}\) When some Case is assigned \textit{in situ} to this internal argument, it is the Partitive, the so-called inherent Case.

\(^{121}\) In English, for instance, the preposition \textit{for} is also an Accusative Case assigner when it is a complementizer; \textit{for} is a preposition in other contexts, though; hence it is not clear whether this property of \textit{for} as complementizer is related to its nature as a preposition. This fact will be relevant for when some contexts with Capeverdean \textit{pa} (a preposition and also a complementizer) are addressed, in Chapter Six.

\(^{122}\) Burzio’s Generalization: All and only the verbs that can assign a \( \theta \)-role to the subject can assign accusative Case to an object (Burzio 1986:178).

Sigurðsson 2006 raises some questions regarding this conspiracy between theta-theory (semantics) and Case theory (syntax/morphology) and proposes a different formulation for the principle:
- All and only the verbs that take a Nom subject can assign structural Acc to an object (Sigurðsson 2006:290).

Mainly inspired by Icelandic data, this author proposes, under the minimalist feature approach to syntax, that Nom is the first Case to be merged; this Low Nominative Hypothesis (this Case is merged low in the structure and then, after Acc has also been merged, raises past to it to a higher position) is presented as a solution to the so-called Nominative Puzzle: if Acc were assigned first (as is traditionally accepted, since the derivation operates bottom to top), how come it has ‘information’ about what will happen with Nom? In other words, if Acc is the dependent Case (and assuming that it was assigned first), how is it aware of the presence/absence of the Case it is dependent on? See Sigurðsson 2006 for more details on this topic.

\(^{123}\) The other traditional reason for this internal argument to move is the EPP. Either way – because of Case or the EPP, or both (this feeds the redundancy arguments against these triggers for movement) –, unaccusatives take an appearance similar to unergatives (like ‘phone’), which take only the external argument.
Also for this internal Case MP has brought a novel, unifying, proposal: in order to be assigned Accusative Case, a DP object must move from its base complement position (sister of V) which is said not to be in the checking domain of V.

In (161a) we have a VP-shell for a transitive structure, which includes the functional projection little $\upsilon P$. This structure is different from the one for unergative verbs (the traditional intransitives, such as ‘run’, or ‘dance’, which sometimes assign Case – ‘run a mile’, ‘dance a waltz’ –, and sometimes do not), in (161b) and also the one for the unaccusatives (now called intransitives), in (161c).

(161)  a. VP-shell for transitives

```
  uP
     /\   \\
   DP1 /   \ u'
        / \   \\
       \  v  VP
            /   \\
           DP2
```

b. Intransitive unergatives

```
  uP
     /\   \\
   DP /   \ u'
      /     \\
     \  v  \\
```

c. Intransitive unaccusatives:

```
  VP
      /\   \\
    V  DP
```

Recall that V has no Case feature; Accusative Case feature is on $\upsilon$ and Nominative Case feature is on Infl.
First of all, we must underline here that this proposal seems to support Burzio’s Generalization. The reasoning goes as follows.

Since little $\nu$:

i) has an external theta-role (say, Agent) to be assigned to its specifier position (see footnote 122, for Burzio’s Generalization);

and

ii) selects as its complement the maximal projection of a verb that, on its turn, only assigns one, internal, theta-role (say, Theme);

therefore little $\nu$:

iii) is the locus of an Accusative Case feature$^{124}$.

According to this view, only $\nu$ has this verb dependent property – Accusative Case feature –, which “belongs” to the DP complement of the verb, not to the verb itself.

Contrasting with a transitive verb, and following Hale & Kaiser 1993, an unaccusative verb is ‘intransitive’, for it assigns no Accusative Case. The traditional intransitive verbs are called ‘unergatives’, because of ergative languages, in which the subjects of transitive verbs are marked ergative case and the subjects of traditional intransitive verbs are marked absolutive – like the objects of transitive verbs$^{125}$. For the sake of clarity, I will adopt a transitive-unaccusative-unergative distinction whenever the argumental structure of verbs is an issue, and avoid the term intransitive, which might cause some confusion.

Resuming the tree in (161a): in order to check Accusative Case the DP$_2$ moves to a higher Spec of $\nu$ – under this approach, a head can have multiple specifiers, accommodating multiple subjects in some languages.

$^{124}$ Notice that the “conspiracy” between the theta-theory and the Case theory, denounced in Sigurðsson 2006, is taken further in this minimalist proposal.

$^{125}$ In other words: absolutive case is assigned to the direct object, if the clause has one; otherwise it is assigned to the subject, given this is the subject of a finite verb; ergative case is assigned to the subject of a finite verb if the sentence has a direct object. Some languages display split-ergativity: both patterns – ergative/absolutive and nominative/accusative – coexist, each of them being at issue in different tenses.
There are no Relativized Minimality effects (there is no restriction on its crossing of DP1), since they are in the same minimal domain. Following this step, DP2 does not move to Spec,IP, a Case position, since it has already lost its Case feature – this [-interpretable] feature has already been checked and deleted, and possibly erased.\textsuperscript{126}

Notice, however, that the type of local relation that is necessary for checking to take place has been questioned by some authors. In Bobaljik 1995, for instance, it is defended that also a head-complement relation is a potential checking relation, as well as a specifier-head relation.

To sum up, there are plenty of open issues concerning assignment/checking of Accusative Case. This short summary, I hope, has shown what is mainly at stake in this discussion around the syntax of this dependent Case.

One of the goals of MP concerning Accusative Case has been to achieve an approach that unifies the requirements and relations involved both in Accusative and Nominative.

There are some critical views condemning these strictly syntactic proposals on Case and agreement. With the help of the label “abstract” (abstract Case and also agreement as a relation between abstract features), these syntactic proposals are missing the main issue, at least with respect to agreement (as pointed in Bobaljik 2006): it is a matter of morphology. This is the topic in the next point.

### 3.1.1.3 A critical view: agreement is not abstract

There are some lines of research conducted by different authors (for instance Marantz 1991 and subsequent work, with ideas also developed in Harley 1995, McFadden 2004, a.o.), presenting case as a morphological matter:

i) it occurs after syntax;

ii) although its input is syntactic, it has no consequences on syntactic operations (as opposed to the abstract conceptions of Case).

\textsuperscript{126} Multiple feature checking means that a DP may have its Case features checked at one position and the phi-features or EPP-features checked at a different position – this also accounts for ergativity.
According to these studies, there are three types of case:

- lexical case (includes quirky case, which is assigned by particular lexical items, such as verbs);
- default case (nominative for nominative-accusative languages; absolutive for ergative languages);
- dependent case (assigned only when more than one NP in a domain may receive case – in nominative-accusative languages, this one is accusative).

The algorithm for assigning morphological case must, therefore, take into account the syntactic structure (such as the decisive hierarchical positions between the NPs that will be assigned nominative and accusative, for instance).

Once this syntactic input is available, the morphological component of the grammar proceeds with a set of operations that will result in the phonological form.

These operations are invisible at the syntax-semantics interface.

Consider, again, quirky case in Icelandic – what we hear is a marker for dative on DPs that behave syntactically as subjects (as their nominative counterparts); hence, this case must be completely distinct from the abstract Case as a licensing condition.

As for agreement, for the present purposes I have chosen to concentrate in summarizing the arguments in Bobaljik 2006, which sustain the view that agreement, too, is a post-syntactic operation, which takes syntactic material and also morphological case as its input (we shall see below why the author defends that agreement occurs post-morphological case) and, thus, has effects only at the PF interface (to which syntax is blind).

The author shows that there are, in fact, contexts where φ-features (person, number, gender) are uninterpreted, in the sense that they do not reach the semantic interface. Thus, they can be ignored in resolving some ellipsis contexts.

Consider, for instance, agreement features on verbs in languages such as Brazilian Portuguese (in sentences where the finite verb has been elided).

\[(162) \textit{Ele sempre comprava aqui, mas nós não.} \quad [\text{comprávamos}]\]

3SG always buy.PAST.3SG here, but 1PL NEG. buy.PAST.1PL

‘He used to buy things here, but we didn’t.’
This is different from stating that $\phi$-features are interpreted on nouns but not on verbs (as the standard minimalist approach on agreement proposes, predicting that this is a syntactic phenomenon).

Following a suggestion attributed to Irene Heim, Bobaljik 2006 shows that $\phi$-features can be uninterpreted on nominal categories as well. Like in (163), where the sloppy reading is the one available:

(163) I finished my homework and you did [$_{VP}$ $\emptyset$]too. [$_{VP}$ finished your homework]

The absent pronoun ‘your’ is thus analyzed as a bound variable pronoun.

Alternatively to the syntactic view, the author argues that agreement (seen as a feature copying/sharing) occurs as a part of Spell-Out (in the PF branch); therefore, $\phi$-features on the target of agreement (the target corresponds roughly to what is called a ‘goal’ in MP terminology) are not seen by the semantics interface (LF branch) because they are not present in the syntax.

Hence, the controller of agreement (roughly what is called a ‘probe’ in MP terminology), more specifically the NP with which the finite verb agrees, is determined in the following terms: the finite verb agrees with the highest accessible NP in its domain (the core notions are ‘highest’, ‘accessible' and ‘domain’). For more details on this argumentation, see Bobaljik 2006.

The idea essential to my purposes here is: (contrary to Chomsky 2001 characterization of agreement) just as NP licensing turned out not to be about case (taking case as the observable phenomenon with that name – morphological case), there is likewise no interaction between agreement and feature-checking/licensing relations.

Crucially, the interesting interaction that exists is between morphological case and agreement (this one occurs post-case; since case operates at the morphological level, agreement must do it as well).

This is the obvious conclusion from the fact that, whenever there is a non-nominative subject and a nominative non-subject in the same sentence, the verb must agree with the nominative DP and not with the subject. Hence, what counts is the
morphological case – and not the grammatical function – for determining the controller of agreement.

This is visible in the Icelandic sentence in (164) – from Bobaljik 2006 –, where the dative subject is an experiencer that, as shown in the literature on Icelandic quirky case, passes all tests of subjecthood\(^\text{127}\), just as the nominative (default case) object passes all tests of objecthood\(^\text{128}\).

(164)  
\[
\text{Jóni} \quad \text{líkuðu} \quad \text{pessir} \quad \text{sokkar}.
\]
 Jon.DAT like.PL these socks.NOM

‘John likes these socks.’

The finite verb agrees (in number) with the nominative object, and not with the dative subject. This leads to the conclusion that agreement takes place post-case assignment.

The same type of controller determination (where the morphological case wins over the subject function) occurs in ergative languages.

Given these facts, thus, locality and accessibility are both necessary and sufficient conditions for agreement. These conditions are provided by syntactic structure, but the result of agreement is not visible to syntax.

Hence, agreement cannot be taken in terms of licensing or feature-checking.

As for the consequences for the standard minimalist approach, while the probe-goal mechanism might be adequate for syntactic licensing (regulating, in some way that is still obscure, the distribution of NPs), it does not describe the relation expressed in morphological agreement. In conclusion, the motivations that lie underneath the distribution of NPs must be investigated elsewhere.

One final note on this is that this conception of agreement has inspired Miyagawa 2006 to propose a separation of the minimalist terms Agree and valuation. The first still operates in narrow syntax (as minimalism predicted for both), but the latter (valuation

\(^{127}\) As said earlier, quirky subjects, like the ones with some psych-verbs in Icelandic and in Japanese, present structural properties of subjects – such as having the ability to bind reflexive anaphors, control the null subject in a non-finite adjunct clause and land on Spec,IP.

of the probe by the goal) is proposed by this author to be a morpho-syntactic operation, occurring at PF. This one is, thus, conditioned by adjacency, and it is this adjacency requirement that drives agreement-related movement in natural language. I shall return, partly, to this line of reasoning at subsection 3.1.3, when I present my proposal on Capeverdean.

In the three points of this subsection I have addressed the main questions concerning Case as a requirement for licensing DPs, especially subject DPs (Nominative Case).

In 3.1.1.1, I have presented a description of the evolution, in the last decades, of the notions of Case and agreement, and the stipulation of an abstract Case as a resort to account for languages, such as Japanese and Icelandic, which display subjects of finite clauses (hence, assumed as Nominative) morphologically marked for dative (the so-called quirky case, which does not count for licensing/Case Filter).

I have shown that the minimalist proposal of abstract features and the notion that agreement is a relation between uninterpretable and interpretable instances of the same feature results similar to the old notion of government. After all, we have to question why we need an abstract Case in the first place.

In 3.1.1.2 I have summarized some questions involved in Accusative Case assignment; although this is not that relevant for the present work, this summary is needed in order to figure out a bigger picture of what is at stake in the current debates.

Finally, in 3.1.1.3 I have presented one critical view (the main lines in Bobaljik 2006) on Nominative Case and agreement as being related to grammatical phenomena, such as NP licensing; here, the main argument against defining agreement in terms of abstract features is that agreement, just as case, is a morphological matter; hence, it occurs after syntax, at the PF interface, although its input comes from the syntax proper.

A way out of this is to define still more clearly the dividing line between morphological case/morphological agreement (the ones that we can see), on one side, and abstract Case/abstract phi-features on the other. On the other hand, though, and considering all this, it is not clear why we should go on trying to understand syntactic licensing in these problematic terms.

Consider, following these lines of inquiry, that: a) only morphological case and agreement are relevant, and there are doubts on whether and how abstract case and agreement trigger movement; b) no one of these morphological phenomena seems
directly involved in the syntactic licensing of DPs (since this latter occurs at syntax, and
the former occur after syntax). Hence, why is it that, for instance, the Capeverdean
sentence * Ka ta Djon papia ingles is ill-formed, whereas Djon ka ta papia ingles is
not?

For all one knows, there may be no other way out of this. Trying to account for
this contrast in the end of subsection 3.1.2, where Capeverdean facts will be analyzed, I
will make use of the notions and discussions presented in the whole section, and
introduce my proposal for the licensing of Capeverdean subjects. This will be achieved
by way of establishing relations between syntax and the morpho-phonology interface.

Before that, however, in the next subsection (3.1.2) I will present some practical (and
some more theoretical) details involved in the Case licensing for external arguments, in
studies by Wurmbrand, on Icelandic and German, and by Costa, on European
Portuguese. A crucial discussion on the role of the EPP comes in the end of the
subsection.

3.1.2 Subjects and Nominative Case

In this subsection, some further details involved in the minimalist machinery around DP
licensing (including abstract Case and Agree) are discussed. The aim of this discussion
is twofold:

a) to understand the main lines in these minimalist tools;

b) to suggest that they pose some practical problems difficult to circumvent.

An ultimate goal is the following: to keep to the minimalist guiding lines of economy,
clarity, and explanatory adequacy, even though the way to reach them does not always
seem facilitated by the abstract notions at stake here.

For reasons of clarity, they are brought into discussion through the proposals in
Wurmbrand 2001 and Costa 2004. Both these authors contend that the EPP is not
universal. The arguments that sustain the more radical view in Epstein & Seely 2006
(where the very existence of the EPP is questioned or, rather, where it is negated) are
summarized as well.
Wurmbrand 2001 shows, based on empirical evidence from German and Icelandic, that some VP-internal Nominative arguments never move to Spec,TP / IP. One of the relevant examples is the one in (165). The author presents this in an embedded clause in order to avoid the verb-second effects in German matrix clauses.

\[(165) \text{weil noch nie [} \text{VP einer Frau ein Orden verliehen]} \text{wurde}
\]

\[\text{since yet never [} \text{VP a.DAT woman a.NOM medal awarded]} \text{was}
\]

‘…since a woman has never been awarded a medal.’

In such environments Case licensing is established under the government-like Agree configuration, just as in Chomsky 2000. Recall this particular definition of Agree:

i) An unvalued feature F (a probe) on a head H scans its c-command domain for another instance of F (a goal) with which to agree.

ii) If the goal has a value, its value is assigned as the value of the probe.

At this point, an additional side note is necessary. In 2001 “Derivation by Phase”, Chomsky developed a biconditional establishing a direct relation between valuation and interpretability of features. This biconditional, which takes valuation as a lexical encoding of interpretability, follows from the fact that syntax could not inspect a feature and determine whether the semantics will or will not assign an interpretation to it, but could inspect the feature and determine whether it is valued or not (Epstein et al 1998).

This biconditional has later been disputed in other minimalist studies. But, before entering into this, let us figure out how the mechanism would work.

Consider that structural Case is an uninterpretable (u) T feature on a nominal (thus, N would be the probe), in the need for an Agree relation with the interpretable T feature on T (goal); here, the scanning would not operate in the c-commanding domain of the probe (in its base position, N does not c-command T). Thus, relating Nominative Case to an uninterpretable feature on T – uninterpretable φ-features would do – would make T the probe. T may scan its c-commanding domain searching for an N (the goal) with which to agree. N – or DP, since the truly attracted constituent is in this case a
maximal projection\(^{129}\) – is then attracted to the position that allows for the Case feature on T to be checked – this position is the specifier of T. Again, we have a specifier-head relation, as in the GB era.

Either way, it seems the uninterpretable \(\varphi\)-features (person, number, gender) on T are a resort in order to make T the probe and the DP the goal – and justify the movement of the DP to Spec,TP where, in fact, the DP checks its own uninterpretable features (the original \(u\) T feature).

In other words, it is not clear whose interest is guiding movement: the interest of T or the interest of the DP. We may admit that this is in the interest of both, and understand this mutual relation as a form of agreement.\(^{130}\) But here, again – and despite these abstract features are considered as morphological –, we are not surely talking about morphological agreement, in the sense of overt agreement morphology, visible at the PF interface.

Here we are back to the biconditional: it has then been strategically proposed that the uninterpretable \(\varphi\)-features on T, being checked under an Agree relation with the interpretable \(\varphi\)-features on DP / N, would allow for a valuation of the uninterpretable T-feature on the DP.

Pesetsky & Torrego, however, have brought up a separation between valuation (syntax) and interpretability (semantics). In other words, this approach includes two novel combinations of feature properties: uninterpretable/valued and interpretable/unvalued. The T-features on V, for instance, are uninterpretable but valued, whereas the T-features on T are interpretable but unvalued: the T-features on T become valued when T enters an agreement relation with the finite V. Given the limitations of the present work, I will not enter into all the consequences of this analysis by Pesetsky & Torrego.

\(^{129}\) In order to understand which constituent is attracted in which circumstances, we have a head movement generalization:

Suppose a head H attracts a feature of XP as a part of a movement operation,

(i) If XP is the complement of H, copy the head of XP into the local domain of H.
(ii) Otherwise, copy XP into the local domain of H.

Thus, for instance, when T attracts a feature of VP, we have (i); when T attracts a feature of the DP-subject, we have (ii); in this case, the XP (DP) moves to the specifier of H (T).

\(^{130}\) For an extensive discussion on whose interest movement takes place, see Lasnik 1999.
Returning to the arguments in Wurmbrand 2001 on Nominative Case licensing without movement, this author discards all the other previous proposals on the type of data shown in (165). These previous proposals have been:

i) An alleged covert movement of the Nominative argument to the canonical Spec,TP position (this approach maintains that Case licensing occurs under a specifier-head relation);

ii) A representation that posits the Nominative argument on Spec,TP, with all the material to its left being in adjunction to TP;

iii) the resort to a null expletive (*pro*), which, from its Spec,TP position, would check Nominative and agreement features and transfer these to the VP-internal argument (in situ) via co-indexation.\(^{131}\)

Against these accounts, the author shows that:

i) The relevant Nominative arguments cannot be in Spec,TP;

ii) Covert movement of the relevant Nominative DP is excluded in certain cases;

iii) The postulation of an expletive *pro* is neither necessary nor motivated.

Crucially, she defends that German and Icelandic lack any syntactic or phonological requirement that Spec,TP be filled, despite having TP and related functional projections. In other words, as the author contends, German and Icelandic lack EPP effects, whereas English and West Flemish are EPP languages.

We must therefore admit that EPP is not universal. Alternatively, it is subject to parametric variation\(^{132}\). This allows for the DP subject in sentences like the one in (165) to be licensed without movement to Spec,TP. This may be so for two reasons: a) the DP needs not move for any features to be checked (they are checked under a government-like Agree relation), b) nor must Spec,TP be filled.\(^{133}\)

---

\(^{131}\) For the relevant data and synthesis of these approaches, see Wurmbrand 2001 and references therein.

\(^{132}\) Wurmbrand 2001 also shows that there is no correlation between the directionality of a language (head-initial or head-final) and the existence of TP-expletives and the EPP.

\(^{133}\) Given the descriptions above, we might say that the redundancy noticed by different authors on the supposed motivations for the DP to move to Spec,TP is, in fact, more complicated than satisfying both Case and the EPP; recall that, as said earlier, checking Case has in itself a mixed function of checking the uninterpretable phi-features on T and the uninterpretable T feature on the DP. Hence, in the cases where the DP does in fact move to Spec,TP, we are not sure about what exactly motivated it: this feature checking centered on T and the DP, or the EPP.
In Costa 2004, evidence from European Portuguese is presented in favor of the possible checking of Case features under Move or Agree. In contexts where the whole sentence is focused, Move takes precedence over Agree. However, an Agree relation is at stake in contexts where the subjects are focused and the subject-verb inversion obtains, in compliance with requirements imposed by the mapping with prosody. We can see this from the sentences in (166), where we have a subject in Spec,VP of an embedded non-finite clause:

(166)  

a. Decidiram ler todos os alunos esse livro.  
decided-3PL read all the students that book  
b. Querem ler todos os alunos esse livro.  
want-3PL read all the students that book  

The empty subject of the matrix verb is co-referential with the subject of the embedded non-finite verb; the fact that no principle-C effects arise, according to the author, follows from the proposal that, in inversion contexts, (the embedded) Spec,IP is empty (again, we must admit that the EPP is not universal).

This type of in-situ subjects is only possible if the matrix verb does not select a CP as its complement (Costa 2004:97), since an intervening CP would preclude the locality conditions necessary for an Agree relation to obtain.

The author confirms these locality effects with the contrasts visible in certain modal environments: when there are transparency effects (verifiable by the possible clitic climbing, for instance; see Gonçalves 1999), this means that no embedded CP is selected by the matrix verb, which confirms the hypothesis in Bošković 1997 that not all non-finite complements project CP (the absence/presence of an intervening CP is, as expected, determined by the selectional properties of the matrix predicate).

In these circumstances (where no CP intervenes in these inversion contexts) – in other words, when there is no strong phase boundary between the matrix T and the specifier of the embedded VP –, an Agree relation may be established.

Alternatively, and still according to Costa 2004, locality conditions are met if we assume that the subject is stranded on its base position, the specifier of VP, meaning
that the relevant functional head (T) can probe this position, which is at the edge of a strong phase (υP), checking the relevant features.

This means that, although subject-verb inversion in European Portuguese is a matter of discourse information (these subjects are focused), this possibility is primarily conditioned by syntax, since it reflects the possibility of checking Case-features under Move or Agree.

Notice that allowing for this feature checking in situ entails one of two possible solutions with respect to the requirements on Spec,TP:

i) either some expletive pro is inserted there in order to satisfy the EPP (however, this hypothesis is not even realistic, since in this case there would be no feature checking in situ, but rather it would be the expletive pro that would check these features – Case and agreement – and transfer these to the VP-internal argument via co-indexation); or

ii) this position is empty, and EPP is not universal.

Epstein & Seely 2006 go further on the EPP question: their proposal is that EPP does not exist. Under this more radical view, however, English overt expletives – the quasi-argumental ‘it’134 and the existential ‘there’ – seem rather hard to explain.

In the case that:

i) the EPP is not there in order to trigger movement, and

ii) feature checking can occur in situ;

it is not clear why the following sentence in English is ill-formed:

(167) * Will be a man outside.

The authors refuse the EPP as the reason for this, since their main claim is that the EPP be eliminated.

As a resort to this puzzle, they make an interesting proposal: the dissociation between agreement and Case (Epstein & Seely 2006:194). This implies an option between two documents by Chomsky. In other words, it means to be back to the

---

134 One of the traditional arguments in favor of the quasi-argumental nature of ‘it’ is that it is assumed as the controller of the embedded PRO in constructions like ‘It is raining without PRO snowing’. For me it is not clear how a category with no semantic content can be the controller (and, thus, co-referent) of an empty category which, in this specific case, also has no semantic content.
proposal in “Categories and Transformations”, within the *Minimalist Program* (1995), instead of taking the view in “Minimalist Inquiries” (2000) – this one predicts that Case-checking freely occurs with agreement-checking (recall that this is the one taken as a theoretical tool in Wurmbrand 2001).

The new proposal goes as follows (Epstein & Seely 2006:194) (the authors’ text is underlined; the notes in parentheses are my own):

i.) **Agreement is checked under the Probe-Goal analysis**
(maintaining in situ the relevant DP, whose interpretable φ-features match with the uninterpretable φ-features on T; thus, T is the Probe and DP is the Goal; the DP is c-commanded by T; they Match; they Agree; the feature is checked and deleted);

ii.) **Case is checked under spec-head** (the Probe-Goal analysis is not useful here, since the c-commanding element is the one who has the instance of Case-feature that is interpretable – although unvalued; Epstein & Seely 2006 propose an alternative structural relation to the traditional m-command: derivational sisterhood, which obtains between elements x and y when x c-commands y before movement and y c-commands x after movement).

But if the facts above solve the problem of (167), what to say about the construction in (168)?

(168) There will be a man outside.

The authors argue that the Case feature of T is checked by *there*, which raises to Spec, TP according to Moro 1997 (Epstein & Seely 2006:195), and I assume that – just as they propose for the ungrammatical *Will be a man outside, in (167) – the Case of *a man is checked optionally by *be (Partitive Case).

We are, however, back to the big puzzle: what element needing its feature to be checked has the priority?

According to this analysis, it seems that the Case feature on T is the most relevant. But why is that? First of all: what exactly is the Case feature of T?
On the other hand, if it were a matter of coupling the interests of T (which still needs its T feature to get valued) and the ones of the DP (who needs its T feature to get interpretable and valued), why is that, instead of (168), we do not always obtain (169), which would be much more coherent with economy?

(169) A man will be outside.

Epstein & Seely present the sentence in (169) as a crucial argument in favor both of the elimination of EPP and the dissociation of Agreement and Case: “if the Case feature of *a man* and the agreement features of T are valued in situ (under Probe-Goal) then, in fact, movement of *a man* is impossible.’ (Epstein & Seely 2006:193)

As an additional side note, in Pesetsky & Torrego 2005 Nominative Case is defined as unvalued uninterpretable T feature on D.

In the system proposed by these authors, EPP is no longer a simple determination on the subject position (either in the old form, according to which all sentences require a subject, or in the more recent form that required that Spec,TP be filled – as we have seen, some materials other than subjects would do), but rather a property of a feature F. This feature F may be, for instance, an unvalued uninterpretable T feature on C.

They use these definitions in their analysis of English T-to-C, establishing a connection between this head-to-head movement and the presence of this unvalued uninterpretable T feature marked [+EPP] on C. Either the nominative *wh*-phrase moving to Spec,CP or T moving to C (hence, either XP movement or head movement) can satisfy this EPP property of the unvalued uninterpretable T feature on C.

The T feature on T being interpretable but unvalued, when T probes D (N) and establishes an Agree relation with it, the T feature on D is rendered interpretable, but how does it ever get valued? Pesetsky & Torrego 2005 propose that both these unvalued T features – on T and on D – become valued by establishing an Agree relation with the valued T feature on little ν; this T feature on little ν, by the way, is rendered interpretable in the process.

Epstein & Seely 2006 make a brief reference to this definition of the EPP as well, but they discard it by arguing that this does not overcome the ban on moving *a*
man in sentences like (169). Nevertheless, it is not clear how their analysis for a sentence like this does indeed overcome it.

Given all the previous puzzles and discussions, I assume, with Wurmbrand 2001 and Costa 2004, among others, that the EPP is not universal. Furthermore, I assume that it is not active in Capeverdean (we shall see why this is a natural assumption)\(^{135}\).

As for English, unless there is some language-specific requirement that the subject position be filled, the contrast between (167) and (168) is still rather hard to explain.

Given everything we know thus far about these mechanisms:

i) the T-feature on T is interpretable, hence it already satisfies Full Interpretation (and, in doing so, it satisfies LF conditions); nothing else is needed at this level;

ii) the T-feature on DP (a man, in this sentence) is uninterpretable but, as we can see by the word order, nothing has been changed in the structural relation between T and the DP from (167) to (168); hence, this uninterpretable T-feature on the DP (which might be taken as the “real” notion of Nominative Case), in (168), must have been checked in some alternative way, such as, as the authors propose for (167), the Partitive Case assured by the copula – hence, the reason for (167) to be ill-formed must lie elsewhere.

As for any other possible element related to Case, the only thing that the there-insertion/raising would solve would be something about T, but, as said above, T goes very well without this. Unless, as in Pesetsky & Torrego’s system, the T feature on T (in this case, an unvalued interpretable T feature – recall that on C we had an unvalued uninterpretable T feature) is endowed with a [+EPP] property and there on its specifier would satisfy this property.

Obviously, I will leave this question open, and make use of the doubts and discussions around this to concentrate on Capeverdean.

\(^{135}\) This view is also possible on Pesetsky & Torrego’s system: suppose that we take their EPP property of a feature as being a property of a feature on T (besides being a property of an uninterpretable T feature on C); this is compatible with this line of reasoning, as long as this property is not universal either.
What seems reasonable to me now is to suggest that, even though the dissociation between agreement and Case checking seems to provide a path in order to answer some previous questions on English, at least, it does not seem active in German and European Portuguese – one possibility is that it is subject to parametric variation as well, which could be related to these morphological features being strong (German and European Portuguese) or weak (for instance Capeverdean and English). For the moment, however, this is only a speculation. In other words, in German and European Portuguese (as we have just seen), both Nominative Case and agreement have been proposed as being checked under an Agree relation (without movement).

One of the crucial facts about Capeverdean is that we find independent reasons to defend that there is no active EPP either, and yet the DP subject is allowed not to raise to Spec,TP only in certain contexts. Therefore, I will contend that this follows from other language-specific properties.

In this subsection I have presented the arguments by two authors (Wurmbrand 2001, on Icelandic and German, and Costa 2004 on European Portuguese) on the possible feature checking under Agree, in contexts where no material is moved to or inserted in Spec,TP.

Furthermore, I have presented the proposal in Epstein & Seely 2006 that the EPP does not exist, on the basis that the phenomena that have been explained resorting to EPP may be accounted for by different tools. These authors postulate a separation between Case and agreement checking (the former is checked under spec-head, the latter under Agree). Based on Capeverdean data I shall argue, with Wurmbrand 2001 and Costa 2004, that the EPP (although still a mysterious entity) is not universal.

In the next subsection (3.1.3), I hope to articulate a specific proposal for the distribution of DPs in Capeverdean, given all the discussions around the relations between Case and agreement (both in the abstract and morphological sense), the licensing of DPs and (briefly) also the possibility of the canonical subject position remaining empty (this is about the EPP, which will be addressed in more detail in Chapter Five).

I will take the post-syntactic approach to agreement and case (as in Bobaljik 2006) as the most adequate, and also make use of the perspective that syntax generates multiple outputs, which are filtered at the interface level, as in Costa 2004. Bobaljik
1995, and Bošković 2001 have proposed this syntax-interface relation considering in particular the interface with phonology (PF), which filters the syntactic outputs through morpho-phonological rules and constraints.

Since the movement operations involved in the creation of these multiple syntactic outputs must be motivated, and given that there is hitherto no more adequate proposal (where the EPP is excluded, for independent reasons), I shall assume that subject DPs are moved out of their base position for some reason somehow associated with active T and the grammatical function of subject. We may call it Nominative Case, as long as we do not ignore how problematic a term this is (its exact definition is hard to establish and it maintains a problematic relation with morphological case). Either way, one may be certain about one fact: in Capeverdean, this feature of active T can be checked under spec-head.

### 3.1.3 Capeverdean arguments: the syntax-phonology interface

The visible distribution of subjects in Capeverdean can be described in some clear lines:

i) the DP subject of a transitive or unergative verb appears obligatorily pre-TMAs and pre-negation; it seems to be in the canonical Spec,TP position;

ii) the subjects of unaccusative verbs are preferred in postverbal position whenever they are indefinite, but behave like the DPs in i) when definite;

iii) subject clitics surface on T, to the right of *dja* and to the left of every other functional morpheme;

iv) as for clitic doubling contexts, some further diagnostics on topicalization must be taken; however, subject clitics being syntactic entities (argumental), absorbing their theta-role on Spec,VP (they are Dº that nevertheless project a DP) and surfacing on T, a DP subject cannot appear on Spec,TP.

We shall see some more arguments on clitics. Now let us consider the constraints on the positions of subject DPs. The relevant examples in (160) are repeated here, as (170).
With these sentences at hand, some core ideas may be summarized:

i) the EPP (as we might understand it now; recall that this is a notion subject to permanent redefinitions; it will be addressed again in 5.3.2) is at least not universal, and it is possibly not active in Capeverdean (we shall see in Chapter Five that there is no reason to postulate such entity as an expletive pro for sentences in (170d), (170f) and (170h)); hence, this is not a valuable trigger for movement of a DP subject to Spec,TP;  

ii) as for the morphological expression of case and/or agreement, Capeverdean displays two crucial properties: the DP subjects are not

---

136 If one entirely agrees with the radical view in Epstein & Seely 2006, one might find the obligatory there-insertion in sentences like ‘There will be a man outside’ still hard to explain.
marked for Nominative\textsuperscript{137}, nor is the verb inflected for person, number or
gender; we are, therefore, left with (abstract) Nominative Case and/or
agreement interpreted as abstract feature checking as a possible trigger
for movement;

iii) as argued in 3.1.1.3, and under the standard minimalist approach, while
the probe-goal mechanism might be adequate for the syntactic licensing
in certain circumstances, it does not describe the relation expressed in
morphological agreement in languages where agreement morphology is
visible;

iv) therefore, the motivations that lie underneath the distribution of DPs
(consider, for instance, the contrasts in (170a)-(170c)) must be
investigated elsewhere;

v) we are left with some abstract feature we might still call Case, not in the
morphological expression of case as examined in Bobaljik 2006, but in
the abstract sense of a mechanism that, still in an obscure way, regulates
the distribution of DP subjects (hence, it is somehow connected with the
grammatical function) and is traditionally considered dependent on
active T; in Capeverdean, this mechanism attracts subjects to T (they
land on Spec,TP or T itself depending on their category status: XP or
Xº); recall that this is different from the Probe-Goal analysis, according
to which features may be checked under Agree (leaving the DPs in their
base position); in order to consider this distinction, we could assume a
separation in the line proposed in Epstein & Seely 2006 (see page 178),

This abstract mechanism, however, and following intuitions based on Capeverdean
evidence, cannot be something exclusively about T itself, since it would predict that in
expletive constructions something should be merged in the position reserved for

\textsuperscript{137} Only the clitics have a distinction between subject/object morphology, and even these may show, in
spite of a subject function, an object form – depending on the phonological environment; like in the
question: \textit{Dja-\textbf{u} bat? ‘Are you leaving?’}
subjects (an overt expletive or an expletive pro). I am defending that this abstract mechanism holds only in environments where there is in fact a subject DP to be attracted to T.

This means a rejection of the old (from GB era) notion that movement would substitute α, the moved constituent, for an empty specifier. In the approach I am assuming, following Chomsky 1995, it is movement that creates the specifier in the first place. From this, we straightforwardly infer that when there is no DP to be moved the specifier is not created (the fact that T has a second edge feature, in current minimalist terminology, means that something may be merged to it, where ‘may’ does not mean ‘must’). This leaves the question of obligatory overt expletives in languages like English to be accounted for by way of a different mechanism, also related to active T but not to subject DPs. This mechanism could possibly be the EPP, which in Capeverdean is obviously not active\textsuperscript{138}.

Consider a transitive verb, such as the transitive entry of papia ‘speak’ (170a), whose DP subject is generated in Spec,VP. As illustrated in the second sentence in (170c), it is forbidden to surface in this position, being moved to a position higher than all the functional material realized in the sentence, namely negation and TMAs.

If we assume that agreement, as abstract feature-checking, provides no independent evidence for A-movement, we might then consider the following:

i) as discussed and shown earlier, Capeverdean TMA morphemes occur in strict adjacency with the verb; we know this fact from some negative evidence (no material, such as adverbs, is allowed in an intermediate position), and from some positive evidence as well (the morphological operation of Lowering is allowed for -ba – T lowers to V –, which implies that this strict adjacency between both is met\textsuperscript{139});

\textsuperscript{138} If we show that the value of an EPP parameter is negative in Capeverdean, this could be an argument for that-trace effects being related to EPP and show that this explains the lack of these effects in Capeverdean; in English the parameter value is positive, and that is why there are that-t effects (besides obligartorily overt expletives). If this is on the right track – the EPP parameter negative in Capeverdean – this could mean that we don’t need to resort to PRO in non-finite constructions. In Chapter Six, we will be back to these cases, where a subject theta-role is at stake.

\textsuperscript{139} This argument about -ba could seem circular, but we shall see in the next chapter that there is independent evidence for the Lowering of -ba.
ii) thus, the DP subject cannot surface on Spec,VP;

iii) Spec,TP is the first position available for the DP subject (a maximal projection) to surface without interfering in this adjacency.

Therefore:

i) in the fashion of the minimalist proposals in Bobaljik 1995, 2002 and Bošković 2001, one may suggest that, following the copy theory of movement, movement occurs in the syntax, creating a chain of occurrences of the moved element; in other words, syntax generates multiple outputs; 140

ii) this movement is triggered by an abstract motivation that must have something to do not only with active T, but also with the very existence of a subject DP able to satisfy this feature in the course of the derivation; the lack of an overt expletive in Spec,TP in expletive constructions would otherwise be hard to explain;

iii) since this relevant feature is abstract (hence, it is indeed not dependent on morphological case marking), there is no inconvenience in assuming it is also related to a given grammatical function (subject);

iv) the multiple syntactic outputs are then filtered at the PF interface: the pronounced copy is the one that does not violate any morpho-phonological constraint – in the present context, the particular constraint at stake is the adjacency requirements mentioned above 141;

v) both the outputs where the DP subject surfaces on Spec,TP and the subject clitic surfaces on T may be the ones chosen; recall that subject clitics are Dº that project a DP, absorbing their theta-role in Spec,VP; in any case, in order to satisfy adjacency, they must not “be in the way”

---

140 Recall that also in the proposal in Costa 2004 on the distribution of adverbs, syntax generates multiple outputs that are filtered at the interface with semantics (only the interpretable outputs are permitted at LF). This is the same scheme, applied to the PF interface; that is, only the outputs that meet certain morpho-phonological requirements are permitted at PF.

141 Bobaljik 1995, 2002 and Bošković 2001 propose that the default is to pronounce the highest copy, as long as this does not violate any interface constraint; under this view, some contexts of inversion, for instance in European Portuguese (where in certain cases the subject of a transitive appears in postverbal position – this has been analyzed, as in Costa 2004, by way of verb movement to T whereas the DP stays in Spec,VP), would be hard to explain. A different hypothesis would be not to consider this default choice of the copies produced, which would leave for European Portuguese, under certain circumstances that must be investigated, the possibility of pronouncing the lowest copy (as long as it does not violate interface constraints). I must leave this question open.
between TMAs and the verb; hence, their lowest copy of all cannot be pronounced either\textsuperscript{142}.

The mechanism proposed here is somewhat different from the one proposed in Miyagawa 2006, since this author claims that there is a purely syntactic operation of Agree, and that valuation (of the probe by the goal) is a morpho-syntactic operation, occurring at PF. This one is, thus, conditioned by adjacency, and this adjacency requirement drives agreement-related movement in natural language.

Now consider other subjects. If this plot is perfectly comprehensible also for the subjects of unergative verbs, some different characters are needed (although applying the same scheme) to account for the subjects of unaccusatives.

These are generated in a postverbal position, where they are assigned their internal theta-role. As noted earlier (and as illustrated in (170d) and (170e)), whenever they are indefinite they are preferred in situ, but they must surface preverbally when they are definite. With respect to the former, and given the scheme above, nothing seems problematic\textsuperscript{143}.

As for the latter, it has been assumed that there must be something about the semantics of definiteness that is at stake here, choosing between both syntactic outputs (blocking the postverbal DP and permitting the preverbal option). Some independent reason at LF must actually be rejecting definite subjects in a postverbal position\textsuperscript{144}.

Recall that there is the following proposal in Costa 2004 for some inversion contexts in European Portuguese (although the author is analyzing transitive contexts): when the whole sentence is focused, Move takes precedence over Agree (and the

\textsuperscript{142} If my proposal is on the right track, it may also account for the specific position of dja. Suppose that, for some further idiosyncrasy about dja, it is not submitted to the same adjacency requirements, as ta and sata, regarding the verb. This might explain why the clitic is allowed on a head to the right of dja.

One different clue to be followed in further research, regarding these morpho-phonological rules and constraints and any Case dependent motivations for movement, is the suggestion in Epstein & Selly 2006:34 that Case may be a PF uninterpretable feature, and not an LF uninterpretable feature.

\textsuperscript{143} Traditionally, and under the rule of the Case Filter, some type of case – inherent Partitive seems the option, in the fashion proposed in Beletti 1988 for postverbal indefinite subjects in Italian passive constructions – is assigned by the verb under a V-complement relation. This author points out that semantically this type of inherent Case, assigned by a verb which does not assign structural Case (Italian passive verbs, or, say, Capeverdean unaccusatives), is only compatible with indefinite NPs.

\textsuperscript{144} In the traditional way, it might be, for instance, the incompatibility between inherent, Partitive, case and definite DPs.

For a discussion on the distribution of indefinites, see Diesing 1992.
subject appears on the specifier of the relevant functional projection). However, an Agree relation is at stake in contexts where the subjects are focused and the subject-verb inversion obtains, in compliance with requirements imposed by the mapping with prosody.

We might also argue, in a parallel with this – that is, with recourse to information structure effects – that there is something about Capeverdean definite subjects which is incompatible with focus.

The obvious consequences of this are the following:

i) the syntactic output which leaves the definite subject of unaccusatives in this base position (being the right edge of the clause a focus position in Capeverdean) is barred at the semantic interface;

ii) the one which presents the definite subject in a position such as the whole sentence is focused (and no morpho-phonological rules and constraints are violated\footnote{I am referring here to sentences like *Ta Djon txiga, for instance.}) is the one chosen;

iii) these positions are, respectively, Spec,TP, when the subject is a DP –\textit{Djon txiga purmeru} ‘Djon arrived first’ – and T, when the subject is a clitic –\textit{E txiga purmeru} ‘He arrived first’.

The similar behavior of definite subjects of unaccusatives, regarding the subjects of transitives and unergatives, must have something to do with their grammatical function. Before proposing that indefinite subjects of unaccusatives are (for some semantic reason) allowed in their base position, we must account for the trigger for movement of their definite counterparts. What we have at the interface is not a trigger for movement, but the application of a filter. Therefore, the trigger for movement, in the syntax proper, must have been of the same nature as the one in transitive and unergative constructions: some abstract feature related both to active T and the DP with the grammatical function of subject. The only difference is that, in unaccusative constructions, the lowest copy of the indefinite DP is allowed at the interface and the lowest copy of a definite DP is not.

In this subsection, I have proposed an alternative perspective to account for the distribution of DPs in Capeverdean. Despite all the mentioned flaws in Case theory (namely, the independency between morphological case and Case viewed as an abstract
feature), and also the disparateness between agreement as an abstract trigger for movement and morphological agreement, I must concede that the movement of subject DPs must be motivated in syntax. I have proposed that this movement is triggered by an abstract motivation connected not only with active T, as in the traditional views on Nominative Case, but also with the very existence of a DP subject able to satisfy this active T feature in the course of the derivation (the lack of an overt expletive in Spec,TP in expletive constructions would otherwise be hard to explain). If one considers that this feature is abstract, there is no contradiction in assuming that it also depends on the grammatical function of the DP. This accounts for the contrast between the requirements imposed on DPs which are not subjects (and thus need not move to Spec,TP) and the ones which are subjects.

I have applied to this the view in some previous minimalist proposals (such as in Bobaljik 1995, 2002 and Bošković 2001) according to which syntax generates multiple outputs, which are, at the phonological interface, filtered by morpho-phonological rules and constraints. In Capeverdean, the constraint at stake is on the adjacency requirements between the functional morphemes and the verb (verified independently – no material is allowed in this intermediate position; the affixing of -ba).

Thus, subject DPs surface on Spec,TP (which is created in the course of this operation) and subject clitics (Xº) surface on T. Morphological case is visible for pronominal forms, although clitics are the only ones that may, under certain phonological conditions, assume an object shape.

As for indefinite subjects of unaccusatives, they surface in their base, postverbal position (assuming the EPP is not active in Capeverdean, Spec,TP is not a problem here, just as it is not a problem when the subject is a clitic, surfacing on T). Definite subjects of unaccusatives are blocked in postverbal position for some semantic constraint related to definiteness (the definite DP may be incompatible either with Partitive Case or with the focus position); they obey the same scheme as the subjects of transitive verbs when they are moved to a preverbal position. Either way, their movement must be dependent on this abstract feature (possibly abstract Nominative Case) dependent on active T and the subject grammatical function.
Before this tentative proposal on the distribution of DPs in Capeverdean, I had to review some theoretical approaches to Case, its relation to agreement and also some minimalist theoretical tools, such as feature checking and Agree (3.1.1).

The questions around the specific licensing of subject DPs also required some notes on the EPP (3.1.2), with the empirical evidence from German (in Wurmbrand 2001) and from European Portuguese (in Costa 2004) showing that the EPP is not universal – contrarily to the more radical proposal in Epstein & Seely 2006, that the EPP does not exist. The EPP will be at issue again in Chapter Five.

In the next section I shall address some verb raising issues.

In subsection 3.2.1 I will summarize some remarks on the motivations for verb raising. In 3.2.2 Capeverdean V-to-T is under discussion, as well as the classical diagnostics for this transformation: the behavior of floating quantifiers (3.2.2.1) – which must be handled with care – and adverbs (3.2.2.2).

Finally, in subsection 3.2.3 I will recall and put into perspective some arguments around verb raising and the affix -ba. In the end, I hope it will finally become clear why I have hitherto been defending that there is no empirical evidence in favor of V-to-T in the language.

### 3.2 Verb movement

Verb movement, as an instance of head-to-head movement, concerns the displacement of the main verbal unit of a clause from the lexical head where it has been generated to some head above it. The head which is the target of the verb can be some upper V, but it can also be a functional head: for instance little υ, or an inflectional head like I or T.

The relevant type of verb movement that I will investigate in this section is verb raising, or V-to-I: verb to inflection. For Capeverdean, and according to what has been proposed in Chapter Two, I will call V-to-T to this hypothesis, since T is the only relevant functional head on the inflectional domain of Capeverdean sentences.

In subsection 3.2.1 I will summarize the motivations for this type of head-to-head movement. In 3.2.2 I will summarize some of the Capeverdean facts that I have presented in previous works, in the light of classical diagnostics for V-to-T in a given
language (even though there are no evident motivations for it – as we will see in 3.2.1, the relation between motivations and V-to-T is not a biconditional). In 3.2.3 I will recall previous arguments on not taking the temporal affix -ba as a proof of V-to-T either. The main conclusion in this section will be that, up to this point in my investigation, there is no clear evidence in order to posit V-to-T in the language.

3.2.1 Motivations for V-to-I

Since Emonds 1978 the distinction between the positions of English finite verbs and French finite verbs regarding VP adverbs, clausal negation and floating quantifiers has been subject to study and debate. Those elements were considered to be generated in the left edge of VP; it was not clear why and how they would appear in a postverbal position in French and in a preverbal position in English.

The hypothesis that in French the verb moves to a position higher than these materials whereas in English it stays in situ has been one of the most developed (there have been others, focusing on what happens to adverbs and floating quantifiers).

Pollock 1989 has proposed, within the Principles and Parameters approach, that there is something about English functional head I involved in this distinction: English I is theta-opaque, which means that it does not allow a verb that assigns theta-roles – that has argument structure – to move there. That is why auxiliaries may raise to I and other verbs may not.

Therefore, this theta opacity or transparency is a property that varies across languages, the value of this parameter being connected with the richness of verbal agreement morphology. Since in earlier stages of English there was verb movement and also one richer agreement morphology, the two properties seem indeed related.

In other words, the crucial parameter involved in verb movement (this one is, then, only a consequence) is this theta transparency of I. Its value is negative in English and positive in French.

In Chomsky 1993 it has been proposed that the relevant parameter is in the value of an abstract morphological feature that licenses verbs and is associated with I. This is I’s V-feature. As other morphological features, these features on V and I must be checked before LF, in order not to violate Full Interpretation. The value for this feature varies
between strong and weak: when it is strong, there is V-to-I movement and the feature is checked and deleted; when it is weak, hence not visible at PF, it will not violate Full Interpretation and, thanks to Procrastinate Principle, its movement and checking is delayed to post-Spell Out.

Besides the great discussion about what one can consider “rich” agreement morphology, there is also the question on the nature of the connectedness between this and V-to-I.

As stated in Roberts 1999, the comparative analysis of diachronic facts in different languages has revealed that there is no biconditional relation between the strength of I’s V-feature and verb movement. In other words, it seems clear that when a language has strong I’s V-feature it has verb movement as well, but when a language loses this strong I’ V-feature this loss does not entail the immediate loss of verb movement. Take Danish, for instance, which has lost the relevant morphology around 1400 and, in spite of that, its verb movement in subordinate clauses survived until the seventeenth century. Or English, in which there has been a gap of some decades between the loss of agreement morphology and the loss of verb movement. Other examples of languages with little or no overt verbal agreement morphology that still allow verb raising are Kronoby Swedish and Tromsø Norwegian (on these Scandinavian languages see Thráinsson 2003 and references therein, particularly Holmberg & Platzack 1995, Vikner 1995, Rohrbacher 1999, a.o.).

The broad debate around “rich” verbal agreement morphology – involving the relevant role of number and person particular distinctions – is not at stake in the present work, for Capeverdean has no verbal agreement morphology whatsoever. This is true for person and number, and applies to any temporal form of any verb.

This, as has been rendered clear, is not a sufficient condition to prevent the language from having V-to-T movement. Baptista 2002 argues that “contrary to the predictions of various V-raising analyses, some creoles like CVC [Capeverdean Creole] display evidence of verb movement in spite of their minimal verbal morphology and the absence of overt subject-agreement morphology.” (Baptista 2002:169)

146 The author considers the Capeverdean verbal morphology as “minimal” – instead of totally non-existent – because of the suffix -ba, which has been described in the previous chapter. However, -ba is a TMA morpheme, and not a person or number agreement morpheme.
The next step involves some widely assumed independent diagnostics in order to figure out if Capeverdean verbs are indeed displaced from their base position and moved up to the inflectional head T. This is the goal of the next subsection.

The main final claim will be that, besides there being no reason to consider that I’s V-feature in Capeverdean is strong (triggering V-to-I), there is also no independent empirical evidence to consider that such a movement takes place in the language.

I will keep away from any discussion concerning the traditional expectations on the negative value for this parameter and the fact that Capeverdean is a Creole language. This would imply considering that the default value for this parameter is negative and also that Creole languages might display some preference for unmarked structures and default parameter values. I leave this debate for more specific creolist studies.147

3.2.2 Classical diagnostics

In some previous works I hope to have been showing that there is no evidence for verb raising in Capeverdean. In this subsection I will resume these previous arguments.

This is a crucial topic, not only because it involves some of the core notions in syntactic theory within Generative Grammar – such as the very nature of parameters and the empirical tools that we may consider in order to study their acquisition by children –, but also because many other linguistic facts in a given language may be analyzed differently according to what the value of this parameter is in the same language.

Therefore, we have to search independent empirical evidence for positing one or the other value. In other words, there are some phenomena that over the last decades have been considered explicable according to which value for this parameter is fixed in the language under study.

The phenomena proposed have always been, of course, a matter of great discussion and successive proposals and retreats. I am presenting here a summary of the argumentation around two of them: the distribution of floating quantifiers (3.2.2.1) and the distribution of adverbs (3.2.2.2).

3.2.2.1 Floating quantifiers

One type of data that has been traditionally used to diagnose verb raising in SVO languages involves the distribution of floating quantifiers. The argumentation (mainly since Sportiche 1988) runs as follows: consider a quantifier modifying a subject NP; whenever it appears to the right of the verb, it must have been left behind by the NP subject moving to Spec,IP; the quantifier is, thus, stranded in Spec,VP (where the NP subject has been base-generated) and the verb raises past it.

The analysis proposed by Bobaljik 1995, however, shows that quantifiers may not be a reliable diagnostics for tracing NP base positions, since they can be analyzed as adjuncts to the left of some XP. This accounts for the following contrast in English:

(171) a. Larry, Darryl and Darryl came into the café *all.
    b. Larry, Darryl and Darryl came into the café all [at the same time].
    c. Larry, Darryl and Darryl came into the café all [very tired].

One could go on (and some researchers do) studying what exactly is the contribution of data with quantifiers (and the question of whether they should be characterized as “floating” in the first place)\textsuperscript{148} for the diagnostics of V-to-T. I will strategically circumvent this topic here, since, no matter which perspective is adopted, Capeverdean data with respect to this are not problematic. Let me show why this is so.

The main data with floating quantifiers provided in Baptista 2002 as evidence for V-to-I involves clauses with unaccusative verbs (like \textit{txiga} ‘arrive’).

(172) a. \textit{Tudu konbidadu txiga na mismu tenpu}.
    ‘All the guests arrived at the same time.’
    b. \textit{Konbidadu txiga tudu na mismu tenpu}.
    ‘The guests arrived all at the same time.’

\textsuperscript{148} I must point that, curiously enough, Bobaljik 2001 and Bošković 2004, both on this topic, have suggestive titles (respectively): “Floating Quantifiers: Handle with care” and “Be careful where you float your quantifiers”.
The author argues that the subject NP here is base-generated in the specifier of V and that the tree representing the sentence in (172b) ‘crucially shows that the verb has moved to Iº, past the quantifier’ (Baptista 2002:197).

One of the reasons why this analysis is problematic is visible in the English translation of the sentence in (172b), which is not ill-formed and exhibits exactly the same word order as Capeverdean. It is widely assumed that English does not have verb movement in the sense under discussion. Hence, it is not clear how this order – quantifiers in a postverbal position – could be a proof for the existence of such a movement when we consider this paraphrase/translation.

Only to illustrate one of the possible different views on this order (other than the one postulating that the quantifier is stranded postverbally) – and also on how the distribution of quantifiers must, indeed, be handled with care –, we have the proposal in Bobaljik 1995, according to which this order can be accounted for in the following way: the quantifier tudu ‘all’ is an adjunct to the left of the PP na mismu tempu ‘at the same time’.

This can be confirmed by the ill-formedness in (173), which goes in a par with the English version in (171a):

(173) * Konbidadu txiga tudu.

Crucially, though, the main reason preventing these sentences from being a diagnostics for V-to-I (V-to-T) movement is as follows: the subject of unaccusative verbs is assumed not to be base-generated in Spec,VP, but rather as the internal argument of the predicate.

Thus, in these constructions, whenever the NP moves to Spec,IP it may very well leave the quantifier behind, even if this is the case that one assumes they are stranded in the previous DP position. This base position is, in these constructions, postverbal. Hence, there is no reason to posit that the verb does not surface on V.

We must conclude that the prediction above – V-to-T inferred from the relative positions of the verb and the quantifier (regardless whether this is correct or not) – is not even applicable in this case.

Therefore, this diagnostics must be confronted with, for instance, transitive verbs; my own informants from Santiago Island consistently reject the quantifier tudu
‘all’ (interpreted as being left by the DP subject in its base position) to the right of a transitive entry. Let us take the contrast in (174a) and (174b). In (174c) we have an example of the quantifier in postverbal position; it has a different interpretation, though:

    children listened all same story

b. Tudu mininus obi mesmu storia.
    all children listened same story
    ‘All the children listened to the same story.’

c. Mininus kome tudu kuskus ku mantega. (*tudu kuskus / tudu mininus)
    children ate all kuskus with butter
    ‘The children ate all the kuskus…’ / * ‘All the children ate…’

In this point I hope to have shown that, even leaving aside all the problems around the real meaning of some patterns in the distribution of quantifiers, the only contexts where quantifiers may appear in a postverbal position in Capeverdean are unaccusative constructions (172). To take quantifiers as marking the path of the DP to which they might be associated (in the line of Sportiche 1988) would not constitute any proof of V-to-T, since in these constructions the subject DP is generated postverbally.

Either way, this diagnostics does not bring any further empirical independent evidence to posit such a positive value for this parameter in Capeverdean.

Let me now proceed with a different classical diagnostics.

### 3.2.2.2 Adverbs

A different possible criterion to diagnose verb raising is the placement of certain adverbs between the verb and its complement. The reason for this to be taken as diagnostics is grounded on the following line of reasoning: since the relevant adverbs are assumed to be left-adjointed to VP, it may be concluded that verb raising is what allows these adverbs to appear to the right of the verb.

However, some of Baptista’s examples regarding adverb placement are taken from literary texts, like the following
(175) *El ta benba pisadu di dinheru ki ta konpensaba materialmenti*

he ASP come loaded of money that ASP compensated materially

tenpu gastadu.  

‘He would show up full of money that would materially compensate for wasted time.’ (T.V. da Silva, 1990:28)

It is well known that, compared to spontaneous speech, literary texts are not the most reliable means for tapping into the grammatical competence of native speakers.

What needs to be taken into consideration is the systematic rejection, on the part of native speakers that are neither bilingual nor writers, of postverbal adverbs. Even examples with the morphologically light *ben* ‘right’, in (177a), are refused by my informants from Santiago. Firstly, these informants never use the word *ben*. Instead they use *dretu*, and always in final position. Compare the following:

(176) a. *João prende ben se lison.*  
John learnt well his lesson

b. *João prende si lison dretu.*  
John learnt his lesson well

Baptista herself has previously shown some concerns with sentences similar to the one in (176a), which she noted as being from her own dialect (from Brava Island) and so they may not be representative.

Moreover, as shown in Loureiro & Pratas 2003, there are in fact some rare light adverbials that are allowed in postverbal position. Consider the examples in (177).

(177) a. *Maria ta obi so es musika li y dipos e ta bai durmi.*  
Maria TMA listen only this music and after 3SG TMA go sleep

‘Maria is going to listen only to this music and then she’ll go to sleep.’

b. *N ta kumeba txeu pexi.*  
1sg TMA eat lot fish

‘I used to eat a lot of fish.’
These, however, have not an adverbial reading, but rather are modifiers of the DP object (which must be quantized, for this position/status of \textit{txeu} to be available). Hence, they may be in left-adjunction to the DP, which does not interfere with the adjacency verb-complement. When we test \textit{txeu} with a clear adverbial reading, we obtain the ill-formededness of the position between the verb and the complement; the sentence-final position is thus the one available, as in (178).

(178) \hspace{1em} a. \textit{N ta ruspetaba nha mai txeu.}  
\hspace{2em} ‘I respected my mother a lot.’  
\hspace{1em} b. \textit{* N ta ruspetaba txeu nha mai.}

Contrast this with Portuguese, where verb raising is widely attested:

(179) \textit{Eu respeitava \textit{muito} a minha mãe.}

These sources cast some doubt on the statement that there is ‘ample evidence’ for verb movement in Capeverdean (Baptista 2002:200). Taking both the floating quantifiers and the postverbal adverbs as hypothetical diagnostics, I would contend precisely the opposite: there is positively no evidence in favor of V-to-T.

In the next subsection we shall finally address what could be taken, at a first glance, as a serious challenge to my proposal, that is, as a serious argument in favor of V-to-T.

\textbf{3.2.3 The -ba problem, again}

Still on verb raising, Baptista 2002 provides a comparison between the clausal architecture of Capeverdean and Guinea-Bissau Creole (GBC). She argues that in GBC ‘there is no evidence for an independent TP, whereas there is in CVC’ (Baptista 2002:204).

This is associated with the fact that the TMA marker -\textit{ba} ‘is a verbal inflection found exclusively bound to verb stems in Capeverdean, whereas \textit{ba} is a non-inflectional (unbound) Tense marker in GBC found not only after verbs, but after nominal and adjectival predicates.’ (Baptista 2002:201) Accordingly, the data are:
Arguing to be following Bobaljik & Thráinsson 1998, on the relation between a Split IP and verb raising, Baptista proposes that this difference between the two Creole languages “predicts that there should be a difference in V-movement between CVC and GBC, and we have shown that this difference does exist” (Baptista 2002:209).

The author contends that the -ba affix can be viewed “as the trigger for the syntactic movement of the verb to T°.” (Baptista 2002:203) True, the author also states that “this is not an uncontroversial issue”, referring to arguments in Koopman 1984 and in Bobaljik 1995 that go against a direct relation between the phonological realization of an affix and the trigger for V-movement. In other words, the necessity for the verb to move in order to check its Tense features is independent of the morpho-phonological content of the affix. The author uses this precisely to argue that this is what happens in Capeverdean, where the lack of suffixation in simple past tense readings for non-stative verbs does not prevent the verb from raising.

Hence, it is not clear how we can take simultaneously the suffix -ba and the lack of other affixes as constituting two arguments (although under different assumptions), both in favor of verb raising.

As I have defended before, and “despite” the only suffix in the language (which I assume cannot be viewed as a trigger), there is no independent empirical evidence to consider V-to-T in Capeverdean.

---

149 Nevertheless, one empirical fact that is used as proof of V-to-T in Capeverdean is also attested in Guinea Bissau Creole: the postverbal position of floating quantifier *tudu*, as in (i) (Baptista 2002:202).

(i) a. *Konbidadu txixa tudu na mismu tenpu.* 
guest arrive all at same time
‘The guests arrived all at the same time.’

b. *Konbidadu tudu txixa na mismu tenpu.* 
guest all arrive at same time
‘The guests all arrived at the same time.’
As for any abstract features involved in the motivations for verb raising, and following minimalist principles of economy and clarity, one could assume that, until finding concluding evidence to the contrary, there are two possible alternative views, both of them without verb raising.

One of them goes as follows: the Capeverdean T’s V-feature is weak, hence not visible at PF, this means it does not violate Full Interpretation and this fact allows its movement and checking to be delayed to post-Spell Out.

The other is rather simple: these features can be checked under Agree.

Now, let me recall briefly (from 2.3.2.1) how we may account for the affixation of -ba without resorting to V-to-T.

This has been inspired by what has already been proposed for the English TMA suffix -ed, since English is assumed – following every other piece of independent evidence – to have no V-to-T either. As I mentioned before, Bobaljik 1995 outlines an analysis for English -ed suffixation, explaining this by means of lowering, a post-syntactic operation, in the morphological component of the grammar.

This is what has been proposed by Costa & Pratas 2003. Following Halle & Marantz 1993 and Embick & Noyer 2001, we have proposed then that postverbal -ba involves lowering of -ba to the verb.

Lowering: “A zero-level element trades its ‘head-to-complement’ relation with its complement for a relation of affixation to the structural head of its complement.”

As we observe in (181).

(181)  
```
TP
  T
   T
    ba V
     VP
      V
```
This alternative view on postverbal -ba makes one interesting prediction: since lowering is a morphological operation, adjacency between -ba and the root is required. It is therefore predicted that there is no material between T and VP, just like it is predicted by the analysis for the English morpheme -ed (Bobaljik 1995). This is confirmed by Capeverdean data.

Furthermore, as mentioned earlier and as we shall see in greater detail in the next chapter, whenever -ba occurs the object clitics are barred, forcing the alternative free pronominal form. As in (182).

(182)  a. Djon odja-l.
       ‘Djon saw him/her.’
    b. * Djon odjaba-l.
    c. Djon odjaba el.
       ‘Djon used to see him/her.’

Under a V-to-T analysis this contrast would be hard to explain. For the pronominal to cliticize there must be strict adjacency; suppose, thus, that it cliticizes before V-to-T, when this adjacency is possible; (182a) is a good result for this. Now, consider that the verb moved taking the clitic with it. How would the grammatical sentence in (182c) obtain? It seems rather tricky.

Alternatively, if we consider the lowering of -ba, the blocking of cliticization, due to the presence of the affix, follows. In which way the presence of the temporal affix precludes the clitic from leaning on the verb is what we will discuss in section 4.1.

In this section I hope to have shown that there is no reason, so far, to posit that the value of the verb raising parameter is positive in Capeverdean.

Considering that:

i) there is no independent empirical evidence – no attested occurrences of the relevant word order variations – to consider that the verb has been displaced from its base position;

ii) there is no overt verbal agreement for person or number in the language; and
iii) the only temporal affix may be accounted for by an operation that does not involve V-to-T;

I propose that the needs of any formal features that require some type of checking under a relation between V and T (either the V features on T or the T features on V) may be satisfied without V-to-T.

3.3 Conclusion

In this chapter, more concentrated on the much discussed role of formal (abstract) features as a trigger for movement (either for A-movement, in the case of DPs, or for head-to-head movement, in the case of verbs), I have defended that:

i) in Capeverdean, most subjects surface on a position higher (say, Spec,TP in the case of DPs, T in the case of clitics) than their base position (except for indefinite subjects of unaccusatives); their movement is triggered by some abstract feature related to the Case properties of T and also some properties of the DP with the subject grammatical function; assuming the copy theory of movement, syntax generates multiple outputs that are filtered at the PF interface; the phonological condition at stake here is the adjacency between the TMAs and the verb;

ii) no matter the abstract V-features on T and/or the abstract T-features on V, there is no empirical evidence so far to posit V-to-T in the language.

Before the conclusion in i), I reviewed some theoretical approaches to Case as involved in the licensing of DPs, its relation to agreement and also some minimalist theoretical tools, such as feature checking and Agree (3.1.1). A critical view has been presented, in particular some arguments from Bobaljik 2006. The questions around the specific licensing of subject DPs also required some notes on the EPP (3.1.2), with the empirical evidence from German (in Wurmbrand 2001) and from European Portuguese (in Costa 2004) showing that EPP is not universal – contrarily to the more radical proposal in Epstein & Seely 2006 that EPP does not exist.
Before the conclusion in ii) I addressed some verb raising issues: in subsection 3.2.1 I summarized some remarks on the motivations for verb raising; in 3.2.2 Capeverdean V-to-T has been under discussion, as well as the classical diagnostics for this transformation: the behavior of floating quantifiers (3.2.2.1) – which must be handled with care – and adverbs (3.2.2.2); finally, in subsection 3.2.3 I have recalled and put into perspective some arguments around verb raising and the affix -ba. In the end, I hope it finally became clear why I have hitherto been defending that there is no empirical evidence in favor of V-to-T in the language.

The next chapter addresses some topics around the verb arguments that are generated in the complement position (the sister of V), that is, the internal arguments. Depending on the selectional properties of the verb, they may bear a subject or an object theta-role.

The realization of object pronominals in different forms, depending on the presence of postverbal -ba (4.1), the absence of an anaphoric expression in certain reflexive contexts (4.2) and the structure of ditransitives (4.3) all have been central topics in different joint works over the last years. The main puzzles will be presented, and some previous and current proposals will be discussed.
Chapter four

Internal arguments

The realization of object pronominals in different forms, depending on the presence of postverbal \(-ba\) (4.1), the absence of an anaphoric expression in certain reflexive contexts (4.2), and the structure of ditransitives (4.3), all may (and they do) have points of intersection and raise doubts and reasonings related to each other. Accounting for these issues involves concepts and tools from phonology and semantics, besides the core questions about structure.

The “object pronominals and \(-ba\)” problem, referred to briefly in Pratas 2002, had been approached in Baptista 2002 as well. There has also been proposed a morpho-syntactic analysis in Costa & Pratas 2003, and a morpho-phonological approach in Pratas 2003b. Thus, the cited work here, Pratas & Salanova 2005, is an accomplishment grounded on the previous ones. The content of section 4.1 is basically the description of this work with Salanova, to which some critics and suggestions have been added.

As for the reflexivity questions (section 4.2), they ended up being the core problem in my Master’s dissertation. Some core problems have been left open, though. The joint work Fiéis & Pratas 2004 makes use of the much relevant proposals in Hornstein 1999, specially the one dispensing with the theta-criterion. The solution we proposed left some questions open as well. In this chapter I present a synthesis of this joint proposal, ending with a different possible way to investigate this topic.

Section 4.3, on the structure of Capeverdean ditransitive constructions, is largely based on a different joint work, Fiéis & Pratas 2005. Also here, in addition to the main points and proposals achieved then, I shall propose some adjustments that seem to me more relevant in the light of the present dissertation as a whole.
Contrarily to 4.1, sections 4.2 and 4.3 will not seem conclusive. The reflexive contexts approached in the former will be addressed again in Chapter Six, since they may bring some insightful information for the analysis of some embedded infinitives. Section 4.4 contains the final remarks on these internal argument issues.

4.1 Objects: clitics vs. free forms

In this section, the earlier mentioned ill-formed co-occurrence of the temporal affix -ba and the object clitic is analysed and a final proposal is made in order to account for this fact. This proposal is mainly based on the analysis of some relevant morpho-phonological rules and constraints that seem active in the language. Again, the necessary perspective involves the syntax interface with morpho-phonology.

First of all, let me recover Capeverdean pronominal forms as in (136), repeated here as (183).

(183) Capeverdean pronominal forms

<table>
<thead>
<tr>
<th>Subject forms</th>
<th>Free forms</th>
<th>Subject clitics</th>
<th>Object clitics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg. ami</td>
<td>mi</td>
<td>N</td>
<td>-m</td>
</tr>
<tr>
<td>2sg (informal) abo</td>
<td>bo</td>
<td>bu</td>
<td>-bu/-u</td>
</tr>
<tr>
<td>2sg (form, masc) anho</td>
<td>nho</td>
<td>nhu</td>
<td></td>
</tr>
<tr>
<td>2sg (form, fem.) anha</td>
<td>nha</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3sg (fem., masc) ael</td>
<td>el</td>
<td>e</td>
<td>-l</td>
</tr>
<tr>
<td>1pl anos</td>
<td>nos</td>
<td>nu</td>
<td>-nu</td>
</tr>
<tr>
<td>2pl anhos</td>
<td>nhos</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3pl aes</td>
<td>es</td>
<td>-s</td>
<td></td>
</tr>
</tbody>
</table>

Let me also recover the general idea (the details involved in any discussion on this have been addressed in Chapter Two) that the bare form of (the so-called) non-stative Capeverdean verbs stands for PastPer and that the form ending in -ba (which has been generated on T and then subject to Lowering, allowing for its affixation to the verb)
generally stands for some type of PastHab or PastPro (depending on the preverbal TMAs, as has also been discussed in Chapter Two).

Having these pieces of information in mind, I describe the central empirical problem of this section in 4.1.1. In 4.1.2 I present the previous approaches to this and in 4.1.3 the main proposal in Pratas & Salanova 2005 will be described – with a few adaptations, namely by excluding here the Optimality Theory instruments included in that paper, since they are not indeed necessary to handle the present phonological facts.

### 4.1.1 The problem

In Capeverdean there is a puzzling incompatibility between object enclitics and verbs affixed with the TMA marker \{-ba\}. The relevant data are summarized in (184):

\[(184) \quad \begin{array}{ll}
a. & N\acute{\text{o}}dja \quad \text{‘I saw’} \\
 b. & N\text{ta}\ o\acute{\text{d}}j\acute{\text{a}}ba \quad \text{‘I used to see’} \\
c. & N\ o\acute{\text{d}}j\acute{\text{a}}-l \quad \text{‘I saw him’} \\
d. & N\ o\acute{\text{d}}j\acute{\text{a}}ba\ e\acute{\text{l}} \quad \text{‘I used to see him’} \\
e. & *N\acute{\text{o}}dj\acute{\text{a}}ba\-l/*o\acute{\text{d}}j\acute{\text{a}}ba\-l/*o\acute{\text{d}}j\acute{\text{a}}ba\-l
\end{array}\]

What, on historical or comparative grounds, could be considered a theme vowel (i.e., the final vowel in \{odja\} ‘see’, \{kre\} ‘want’, \{xinti\} ‘feel’, \{konko\} ‘knock’ and \{bombu\} ‘carry [the baby] on the back’), seems actually part of the stem, for there is no morphological or phonological reasons to think otherwise.

Adding to the empirical puzzle in (184), we know that Capeverdean displays a simple weight-sensitive stress rule that assigns primary stress to the syllable containing the penultimate mora. All coda segments, /r, l, m, n, s/, count as moraic. This rule accounts for the stress pattern of nearly all words:\(^{150}\)

\[(185) \quad \begin{array}{ll}
a. & k\acute{\text{a}}za \quad \text{‘house’}
\end{array}\]

\(^{150}\) Many details could be added, v.g., about the prosodification of rising vs. falling diphthongs, the stress pattern of compounds, etc. These are beside the point here.
In exceptional cases the last consonant has to be marked as extrametrical in the lexicon:

\[(186)\]
\[
\begin{align*}
\text{a. ménos} & \quad \text{‘less’} \\
\text{b. kinhéntus} & \quad \text{‘five hundred’}
\end{align*}
\]

Morphologically complex words – stems+derivational affixes \((187a)\) and \((187b)\) – behave just as expected, including the inflectional suffix \{-s\} ‘plural’, as in \((187c)\), if we assume this one as extrametrical. Moreover, this ‘plural’ \{-s\} is not very productive.

\[(187)\]
\[
\begin{align*}
\text{a. padjínha} & \quad \text{‘weak hashish (pádja)’} \\
\text{b. bistídu} & \quad \text{‘dress (from the V bistí)’} \\
\text{c. kárus} & \quad \text{‘cars’}
\end{align*}
\]

We can check the rarity of this ‘plural’ \{-s\} when we observe the sentences in \((188)\):

\[(188)\]
\[
\begin{align*}
\text{a. Kes} & \quad \text{livru bunitu kusta dinheru txeu.} \\
& \quad \text{these-PL book beautiful cost money lot} \\
& \quad \text{‘These/The beautiful books are expensive.’} \\
\text{b. Tudu mininu gosta di livrus bunitu.} \\
& \quad \text{all child like of book-PL beautiful} \\
& \quad \text{‘Every child likes beautiful books.’}
\end{align*}
\]

In fact, whenever it occurs, this \(/s/\) has to be marked as extrametrical.\(^{151}\)

Slightly more formally, the stress rule in Capeverdean is the following\(^{152}\), illustrated in \((189)\) with the word \{divugádu\} ‘lawyer’.

\(^{151}\) It would also be possible to analyze the special behavior of the plural as following from it belonging to a putative affixal class 2, which wouldn’t affect stress. It has been shown by Ximenes (2004) that the plural morpheme in Brazilian Portuguese is outside of the domain where coda \(/s/\) is a trigger for diphthongization:

\[(i)\]
\[
\begin{align*}
paz & \quad \text{[pajs]} & \quad \text{‘peace’} \\
pá+s & \quad \text{[pas]} & \quad \text{‘shovels’}
\end{align*}
\]

\(^{152}\) The rule here is given loosely within the formalism of Halle & Idsardi (1993) for concreteness.
4.1.2 On previous approaches

Some subsidiary empirical facts, central to previous approaches, must be described here. The first is that, as argued in Baptista 2002, there are no clitic clusters in Capeverdean:

(190) a. * Modi ki el tene kel karu-li?
    How REL 3SG have that car-here?

b. * E si pai ki da-l-l.
    Is his father rel give.3sg.3sg

   ‘It was his father who gave it to him’

c. E si pai ki da-l el
    Is his father rel give-IO3sg DO3sg

   ‘It was his father who gave it to him.’

Encliticization is also impossible after passive verb forms in {-du} (Present) or {-da} (Past). Let us check this also in passives of double object constructions (DOCs).

(191) a. * E odjab-ba-m. / * E flaba-m.
    he see.TMA-1SG / he tell.TMA-1SG

b. * N ta dadu-l. / N ta dádu el.
    I give.TMA-3SG

   ‘I have been given it’
When one of the few free-standing lexical items is allowed postverbally (that is, between the verb and the pronoun)\textsuperscript{153}, encliticization is blocked as well:

(192) a. *N odja so el. / *N odja so-l.
   I see only he
   ‘I saw only him.’

b. *N gosta d’el. / *N gosta di-l.
   I like of he
   ‘I like him.’

Just as it happens when the pronominal is coordinated with another DP, no matter the order between them (with ku having the ‘and’ conjunction reading\textsuperscript{154}):

(193) a. *Nu odja-l ku Maria
   we see.3SG and Maria
   ‘We saw him and Maria.’

b. Nu odja-l ael ku Maria
   we see.3SG 3SG and Maria
   ‘We saw him and Maria.’

c. Nu odja Maria ku el. / *Nu odja Maria ku-l.
   we see Maria with 3SG
   ‘We saw Mary and / with him.’

Notice that the order in (193a) is possible with ku heading a PP (a productive preposition value for this – ‘with’) in right adjunction to VP.

See the examples in (194):

\textsuperscript{153} As described in the previous chapters, there are requirements of strict adjacency between the verb and the arguments:

(i) (Tudu dia) Djon (*tudu dia) ta odja (*tudu dia) tilibison (tudu dia) na ora di djanta.
   (Always) Djon (*always) TMA see (*always) tv (always) in time of dinner.
   ‘Djon always watches tv at dinner.’

\textsuperscript{154} Notice that, as opposed to English ‘and’ and to Portuguese e, Capeverdean ku – which sometimes has an adverbial value, ‘with’ (check (194)) – connects only phrases, not clauses. The word used in coordination contexts is y. We might say that ku has been derived from Portuguese com ‘with’, and in these conjunction values it presents some kind of reanalysis effect; y, on the other hand, has been derived from Portuguese e ‘and’.
(194) Question: *Modi ki nhos ta kume katxupa?*

   how REL 2PL TMA eat katxupa

   ‘How do you eat katxupa?’

Answer: *Nu takume-*l ku kudjer.

   1PL TMA eat.3SG with spoon

   ‘We eat it with a spoon.’

In the previous examples in this subsection we have seen the relevant data that now allow us to confront the approach presented in Baptista 2002 to the empirical puzzle in (184). According to the author, there is a morphological template where TMAs for Past Imperfective or for Passive compete for a slot with person marks. There is V-to-T movement in the language; {-ba} is affixed to the verb in the syntactic component.

In Pratas & Salanova 2005, it has been argued that:

i) the aforementioned competition is a plain stipulation;

ii) putting together direct and indirect object clitics with tense markers is both intuitively and empirically wrong.

Now consider the few irregular imperfective forms in Capeverdean, for instance {era} ‘used to be’, {sabia} ‘used to know’ (alternate with regular {sabeba}) and {tinha} ‘used to have’ (alternate with regular {tenba} or {teneba}):

- we may analyze these as consisting of an allomorph {-a} for Imperfective, which selects for the allomorphs {er-}, {sab-} and {tinh-} of the verbs {e}, {sabe} and {ten/tene}, respectively.155

Therefore we can defend that:

i) there is allomorph selection between the stem and the TMA morpheme, while no such allomorphy occurs between stems and person clitics;

ii) person clitics have corresponding freestanding forms that are required in various contexts such as coordination or focalization, while the TMA ending {-ba} is exclusively bound.

155 Alternatively, one could say that there is suppletion for the V+T. Our point in the mentioned paper was not altered if this alternative analysis is adopted, but our analysis of stress shift and the restriction on clitics depends on a stem being separated from inflectional endings.
In Costa & Pratas 2003 there is a proposal different from the one in Baptista 2002. Following Embick and Noyer 2001, we proposed then that clitics attach to the verbal root, under Leaning, which is defined in (195):

- Leaning (Embick and Noyer 2001):

  “If X is an element peripheral in some constituent C (a.), X will not be able to invert with an element Y that is outside of the constituent C (b.), although leaning is possible. (c.).”
  
  a. [...] * [c X * Z]
  b. [...] X + Y * [c Z] – impossible inversion
  c. [...] Y + X * [c Z] – possible leaning

  a * b means that a must linearly precede b and be adjacent to b.

Since Leaning is a morphological phenomenon, adjacency is a condition for it to occur.

Also crucial in Costa & Pratas 2003 is the consideration of inflection as affixed to the verb not by V-to-T, but rather by lowering of T to V, in the morphology component of the grammar, under strict adjacency. Lowering is constrained in the following way: in a configuration X * [ Y * Z ], where * represents linear precedence and the brackets represent morphological constituency, lowering of X to Y cannot take place, since it would “break up” a morphological constituent.

The consequence of this proposal is the following: if -ba intervenes between the verbal root and the clitic, there is no context for Leaning to apply, which derives the complementary distribution between -ba and the pronominal clitic.

In other words, besides the ungrammaticality of {flábal} there is the following fact: the morphological constituent {fla} + {l} prevents the lowering of {-ba}. Alternatively, if -ba would not lower, this bound morpheme would not constitute a legitimate morphological object.

Independent evidence for Leaning may be found in the fact that there are no clitic clusters, as we have seen in (190) and as we can see in (195). Only one (object) clitic may surface per sentence. This shows that clitics cannot lean onto other clitics:

---

156 This proposal has been sustained over the chapters of this dissertation.
4.1.3 Stress matters

Following Pratas 2003, the argumentation in Pratas & Salanova 2005 favors the view of the ill-formedness of (184e) as being intimately tied to the process of stress shift involved in the encliticization process.

In 4.1.3.1 I will describe the relevant phonological features in Capeverdean, without resorting to an Optimality Theory approach (used in the previous works, Pratas 2003 and Pratas & Salanova 2005), since this is not needed in order to account for the present facts. In 4.1.3.2 I will present the morpho-phonological puzzle in (184) as being the result of a crossing between a rule and a constraint.

4.1.3.1 Some morpho-phonological features in Capeverdean

Let me now recover one crucial phonological property of the language.

Capeverdean weight-sensitive stress rule:

i) Primary stress of a word falls on the syllable containing the penultimate mora.

ii) All coda segments, /r, l, m, n, s/, count as moraic.

This rule accounts for the stress pattern of nearly all words, as we have seen in (185), here repeated as (196):
Morphologically complex words – such as we had in (187a) and (187b), here repeated as (197) – show that the stress shifts to this penultimate mora position when some derivational morpheme is added to the stem:

(197) a. *padjínha ‘weak hashish (pádja)’
    b. *bistídu ‘dress (from the V bisti)’
    c. *kárus ‘cars’

Notice, again, that the {-s} that marks ‘plural’ on nouns and adjectives seems to be marked as extrametrical (197c), whereas, as in (197b) with an adverbial, it counts as moraic as the other coda segments.

Simultaneously, suppose that there is the following (not surprising at all) morphophonological constraint in the language:

- The main stress of a word cannot fall outside the stem.

This is indeed respected in morphologically complex words – stem+derivational affixes.

Now let us check the result of crossing the rule with the constraint on the problem we are dealing with, and see what happens.

### 4.1.3.2 Between a rule and a constraint

The data / puzzle in (184) is here repeated as (198).

(198) a. N ódja ‘I saw’
    b. N ta odjába ‘I used to see’
    c. N odjá-l ‘I saw him’
    d. N odjába el ‘I used to see him’
    e. *N ódjaba-l/*odjába-l/*odjabá-l
The stress rule would require that in \{odja + ba + l\} the stress be placed on the last syllable (where we have two moras). This would violate the morpho-phonological constraint at issue here, which requires stress to fall within the stem.

Thus, \*odja-bá-l is ruled out: if the stress rule of Capeverdean is to be obeyed, stress falls on the tense inflection \{-ba\}, out of the stem.

The fact that the form in (198d) is the one to be chosen follows straightforwardly from the effects of crossing the aforementioned rule and constraint. Being el and -l allomorphs for the same abstract feature, 3sg, this is a simple case of allomorphy determined for morpho-phonological reasons.

For this analysis to hold we must consider:

i) two listed allomorphs for each combination of person and number features (some of these allomorph pairs differ only in being or not being syllabic (like for the 3sg – el / -l – and for the 3pl – es / -s), but the relation of the members in other pairs is not as straightforward);

ii) the free forms as being non-clitics, and as behaving sometimes as XP and sometimes as X° (this last one would be the case here, where they must be heads);

iii) other lexical items (namely the TMA) are already inserted (subject to Lowering) at this point, and there is no reason not to assume this.

There are, however, some potential problems to this analysis.

First of all, we still need to explain why the free form is precluded when the constraint at stake does not hold, that is, when there is no TMA affixed to verb; in other words, when there is the bare form of the verb, like in \* N ódja el. A possible way out of this is to argue that the insertion of the freestanding form of the object pronominal is more costly to the grammar, since it results in the creation of a separate stress domain.

Secondly, one could argue that in Capeverdean there is no other environment where the constraint at issue (stress must fall within the stem) can be confirmed (this seems to be true) and, thus, this would be a circular argumentation. As said before, though, this is not an unexpected constraint cross-linguistically.

Finally, if an object pronominal form follows a word that is not a verb, what prevents it from encliticizing? That is, what prevents the direct object pronominal from encliticizing to an indirect object DP in (199)?
In Pratas & Salanova 2005, an independent constraint has been introduced:
- preventing object pronominals from encliticizing to anything but the verb;
- or, more generally, one according to which a pronominal can only encliticize to something whose syntactic projection the pronominal is inside of.

The latter hypothesis might sound stipulative, but it is parallel to another well known case of prosodically-conditioned allomorphy, the one of the feminine definite article in Spanish. The definite article {la} has the allomorph {el} before words that begin in stressed /a/; this is clearly prompted by hiatus avoidance.

This allomorphy of the definite article does not take place before an adjective, even if a hiatus would result:

Whatever the correct analysis,\textsuperscript{157} this allomorphy is prompted by phonology and sensitive to syntactic information.

Given what we know from Chapter Two, this more general constraint does not hold for subject clitics, which can lean, for instance, on \textit{pa}, \textit{ma} or \textit{dja}, whenever one of these occurs to the left of the clitic.

Still on the specific case in (199b) – the clitic would shift the stress of the word to the penultimate mora and, with / l / counting as moraic –, the phonological result

\textsuperscript{157} Actually, although the alternation itself has received a great deal of attention, we know of no discussion of the fact that it is sensitive to the adjective vs. noun (and even proper noun vs. common noun) distinction.
would be *Da-l-u / *Prizenta-s-u. We see that, to start with, changing the stress of a proper name sounds actually odd. But given that this is a double object context, perhaps there is more to it. This prompts us to the analysis of other ditransitive constructions. These will be addressed in greater detail in the next section.

For the moment, let me just point out what has been considered a challenge to this analysis: two clitics are not permitted even when they could be syllabified as a single mora (IO stands for indirect object and DO for direct object).

(202) a. *Da-l-u / *Prizenta-s-u
    give.3SG (IO) 2SG (DO) / introduce 3SG (IO) 2SG (DO)
b. *Da-l-u
    I give.3SG (IO) 2SG (DO)
    ‘I gave him/her you.’(‘I gave you to him/her.’
b’. *Da-l-u
    I give.2SG (IO) 3SG (DO)
    ‘I gave it/him/her to you.’
c. *Prizenta-s-u
    I introduce.3PL(IO) 2SG (DO)
    ‘I introduced you to them.’

If we followed a plain construction of the hypothetical phonological words in (202a), there would be no problem with the stress constraint at issue here, since the main stress of the word would fall on the verb stem – hence, it would be difficult to explain the prohibition of these words.158

A plausible phonological solution to this could follow if the underlying form of clitics is viewed as containing a mora.159 Forms such as those in (202a) would then get stress outside of the stem, too, and be ruled out for the reasons assumed above.

158 Jonathan Bobaljik, p.c. 2005, has suggested that in this specific context what may be at stake is not so much a phonological phenomenon but rather a morphological constraint which holds in some Romance languages: this is Personal Case Constraint (PCC), which forbids the order V-3sg-2sg in DOCs.

159 We left this as a plain stipulation in Pratas & Salanova 2005, and it will also be left as a stipulation here, since the knowledge of the language phonology is still insufficient to take this as following from other facts (i.e., a general condition on allomorphy). Also, we are aware that there is overlap between the work that is done by positing clitics to be inherently moraic and the stress rule that is proposed for Capeverdean, yet if we have reasons to doubt that coda /s/ is moraic (the number of forms with “extrametrical /s/” far outnumbers that of forms with other extrametrical consonants), the overlap is significantly curtailed.
A much simpler view is that clitics cannot lean on other clitics. Thus, there might be a further (also not surprising at all) constraint as follows:

i) object clitics can only lean on the verb (this would correctly rule out (199b)),
ii) not on the phonological word that contains the verb (this would correctly rule out (202a)).

In this section encliticization in Capeverdean has been analyzed as a case of prosodically conditioned allomorphy. Enclitic forms and free forms are allomorphs that are chosen following the stress rules of the language.

In order to fully analyze the Capeverdean clitic cluster problem we have to consider three sets of stipulations:

i) the morphemes {-ba} and {-du} are clitics, and there are no clitic clusters in the language (Baptista 2002);
ii) V+clitic form a morphological constituent that disallows the lowering of {-ba}, and, in addition, clitics cannot lean on clitics (Costa & Pratas 2003);
iii) stress cannot fall outside of the verb stem, and object clitics can only lean on the verb (and not on the prosodic word that contains the verb).

The latter, which also involves the developments achieved in ii), is the analysis assumed in this dissertation.

The next section deals with a different problem with respect to Capeverdean internal arguments: most of the constructions with a reflexive reading do not realize an overt reflexive pronoun or expression.

There are two possible types of perspectives on this:

i) one mainly syntactically oriented (deriving these structures through syntactic operations such as movement of the “absent” constituent);
ii) the other one mainly semantically oriented (based on argument structure, theta-role selection and lexical operations).
The proposal in Fiéis & Pratas 2004, presented in 4.2.1 makes use of both, and the path leading to it will be described. In the last two years of research, however, new doubts and judgements have motivated a somewhat different view, mixing syntactic and semantic observations, and these will be at issue in 4.2.2.

4.2 Identified moving subjects? Some reflexive contexts

This section addresses a much striking Capeverdean empirical phenomenon, which has been the central focus of my Master’s dissertation: some reflexive contexts do not show any overt reflexive expression.

This may be illustrated with the example in (203), with a similar counterpart in English, only that in Capeverdean these type of construction (Subj.-V sentence, involving a so-called transitive verb\(^{160}\) with a reflexive reading) is much more productive, and is allowed by a great number of predicates.

(203)  *Pedru laba.*

Pedru wash

‘Pedru washed (himself).’

An overt Capeverdean reflexive expression is built with (possessive pronoun)+*kabesa* (like in *si kabesa*, literally ‘his/her head’). The possessive is not obligatory; the version like *Pedru mata kabesa* ‘Pedru killed himself’ is likewise common among my informants, along with *Pedru mata si kabesa*. This overt reflexive expression is only needed (and obligatory) when a reflexive entry of the transitive verb is clearly unusual, or pragmatically odd. Like with *mata* ‘kill’, or *djobe na tilibison* ‘see on tv’, etc.

With a great number of other verbs, the absence of the reflexive is not only allowed; it is indeed preferred.

In subsection 4.2.1 I will present a summary of the main points at stake around this empirical puzzle, the essential consequences of some previous proposals (including my

\(^{160}\) The label “transitive” is meant to refer any verb which selects an internal argument with a direct object type of theta-role (like Theme) and an external argument with a subject-type of theta-role (like Agent or Experiencer). The doubt cast over this label is that, as we can see now, many entries of these verbs do not correspond to this structure. We may say that, apparently, they turn into unaccusative-like entries, or that, as Reinhart 1996 proposes, they suffer a lexical reduction of the internal argument.
own, in Pratas 2002, and the one in the joint work Fiéis & Pratas 2004) and also the problems that have been left open.

In subsection 4.2.2 I will present a complementary approach to this, which by now I believe more logically integrated in the spirit of this dissertation, supported by minimalist concerns of simplicity and adequacy.

### 4.2.1 Plausible analyses

Consider the following constructions with some particular predicates / human subjects:

a. – Capeverdean transitive entries;
b. – Capeverdean reflexive entries;
c. – European Portuguese entries marked by *se*.

(204) a. *Maria laba ropa.*
   Maria wash clothes
b. *Maria laba.*
c. *A Maria lavou-se.*
   ‘Maria washed (herself).’

(205) a. *Zé Luis fri-l*$_{3}$/$_{3}$.
   Zé Luis hurt-3SG
   ‘Zé Luis hurted him/her.’
b. *Zé Luis fri.*
c. *O Zé Luis feriu-se.*
   ‘Zé Luis got hurt.’ / ?? ‘Zé Luis hurted himself (on purpose)’

(206) a. *Djon perde dinheru.*
   ‘Djon lost his/some money.’
b. *Djon perde.*
c. *O Djon perdeu-se.*
   ‘Djon got lost.’ / * ‘Djon lost himself (on purpose)’
Notice that, as we may see by the most direct English translation, only the first Capeverdean example of a reflexive sentence without an overt anaphor – (204b) – has what might be considered an obvious Agent subject. The examples in (205b) and in (206b) show that for some Capeverdean verbs a reflexive environment results in an impossible agenteive reading for the subject.

On the other hand, if we had something like Zé Luis fri si kabesa (di xokota) – on purpose – or Djon perde si kabesa (di xokota), the reading of the anaphoric expression would be, literally, ‘Zé Luís hurt his head (on purpose)’ and ‘Djon lost his head (on purpose)’, a type of anaphoric expression different from what we are discussing here.

In (207) and (208) we have reflexive entries with non-animate subjects (these European Portuguese versions – in c. – also have se as a reflexive marker)\(^{161}\).

\begin{itemize}
\item \textbf{(207)}
  \begin{itemize}
    \item a. Porta abri.
    \item b. \textit{A porta abriu-se}.
  \end{itemize}
  \textit{door open}
  \textit{‘The door opened.’}

\item \textbf{(208)}
  \begin{itemize}
    \item a. Vidru kebra.
    \item b. \textit{O vidro partiu-se}.
  \end{itemize}
  \textit{glass break}
  \textit{‘The glass broke.’}
\end{itemize}

In these contexts, we know that the subject can never have an Agent theta-role. Notice, however, that the English paraphrase does not include the verb ‘get’. What we have, here, is indeed a type of unaccusative entry – the subject has a Theme or Patient theta-role (it has been generated as the verb internal argument).

This organization by types of predicates is of crucial relevance for the proposal I am going to present in subsection 4.2.2. Before that, let me just present the previous proposals, their consequences and the problems they left open.

\(^{161}\) For the transitive entries of these verbs we must have Agent-type of subjects (hence, the DPs in (207) and (208) must be really internal arguments).
4.2.1.1 Where European Portuguese has \textit{se}

According to Levin & Rappaport Hovav 1995, among others, some lexical properties account for different entries of some predicates. This type of variation occurs in a given language, but in a cross-linguistic perspective as well, as shown in Fiéis 2003:

(209) a. \textit{Jean a brûlé la maison}.
   ‘John burnt the house.’

   b. \textit{La maison a brûlé}.
   ‘The house got burnt.’ (literally ‘The house burnt’)

(210) a. \textit{Sortez la voiture du garage}.
   ‘Take the car out of the garage.’ (literally ‘Leave the car of the garage’)

   b. \textit{Jean est sorti}.
   ‘John left.’

(211) a. \textit{Personne ici ne serait capable de me tomber}.
   ‘Nobody here would be able of putting me down.’
   (literally ‘Nobody here would be able of falling me’)

   b. \textit{Jean est tombé}.
   ‘John fell.’

A lexicalist perspective is also adopted by Reinhart (1996, and subsequent work), for whom in some English verbal entries there is a lexical reduction, like in \textit{Max washed} – reduction of the internal argument, resulting in a reflexive reading –, and in \textit{Lucie rolled} – reduction of the external argument, resulting in an unaccusative entry. This proposal has been adopted in Pratas 2002 for the Capeverdean data above.

Some details involved in this proposal are as follows. One crucial notion is the Lexical Uniformity Principle: each verbal concept corresponds to a single lexical entry with a specific thematic structure; the different thematic forms of the same verb are derived through lexical operations. According to Chierchia 1989, reduction applies to a two-place relation; it identifies two arguments and reduces the relation to one property.
The reduction of the internal argument does not erase its theta-role. The verb looks like an intransitive entry. Its structure could be taken as unaccusative; recall, though, that in unaccusatives the argument missing is the external, not the internal.

From the examples above, the cases in (207) and (208) really look like unaccusative entries (they lack the external theta-role), whereas the example in (204) seems to correspond to the ones accounted for by the internal argument reduction proposed in Reinhart 1996 – it has an intransitive structure.

One of the problems of this approach comes when we observe the ones of the type illustrated in (205) and (206), which can be translated into English with what I will call henceforth a ‘get construction’. Whether these can help us have a different perspective on the others – both the supposedly agentive laba ‘wash’ and the supposedly unaccusative entries, such as abre ‘open’ and kebra ‘break’ – will be left for future analyses.

Still concerning Reinhart’s proposal, we may say that in Dutch, just as in Portuguese, this operation of lexical reduction leaves behind a morphological mark of the internal theta-role (and not of the Accusative Case). This mark does not exist in English or Capeverdean.

\[(212) \quad \begin{array}{ll}
  a. \textit{Peter wascht sich}. & \text{(Dutch)} \\
  b. \textit{O Pedro lavou-se}. & \text{(European Portuguese)}
\end{array}\]

The assumption of, for instance, the European Portuguese \textit{se}-clitic as a trace of the Accusative Case is not satisfying, and this is another problem of this approach.

At this point we could try to solve this by assuming that Capeverdean, like English, does not have these morphemes – \textit{se}-type – available in its numeration, hence it can not use them. This does not explain much, though.

In Fiéis & Pratas 2004, it has been proposed that the relevant “visible” contrast between some reflexive entries in different languages – presence-absence of a \textit{se}-morpheme – is determined in the morphological (post-syntactic) module of grammar. Recall that in the
syntax some terminal nodes are generated and moved by syntactic rules and constraints. The result of this constitutes the input of a morpho-phonological domain, the PF branch of the grammar.

We must also recall here that, as assumed earlier in Pratas 2004, the Split,IP parameter has a negative value in Capeverdian, whereas in European Portuguese it has a positive value (IP is split in some other functional maximal projections, such as TP, AgrP, AspP).

Finally, a crucial – although controversial – assumption is the one following the proposal in Hornstein 1999: treating theta-roles (θ-roles) as features on the verb allows a DP to move to a theta-position. This predicts that we can dispense with the theta-criterion, which does not allow a phrase movement from a theta-position to another theta-position. Assuming Horstein's rejection of this constraint, a DP can receive more than one theta-role.

Thus, the line of reasoning in Fiéis & Pratas 2004 goes as follows. In European Portuguese there is a functional head (Refl, for some cases, Caus or Voice, for others) where the se-morpheme is inserted, in the morpho-phonological module. This allows for an unified analysis of the se-clitics – we proposed that all of them are non-argumental, while the previous analyses (Cinque 1988, Duarte 2001, a.o.) considered some of them argumental (like in O Pedro lavou-se ‘Peter washed’) and some non-argumental (A porta abriu-se ‘The door opened’).

In Capeverdian no category other than T is projected in the functional domain, hence there is no place to accommodate this morpheme. This means that the sentence reading depends not only on the semantics of the predicates but on the sentence composition as well (a reflexive reading, for instance, is available whenever there is no expressed internal argument for a transitive verb).

We would have a reflexive reading (213), an inherent reading (214) and an anticausative-unaccusative reading (215) in mono-argumental entries of some Capeverdian predicates (confront with the European Portuguese counterparts, in the b. examples), and these are dependent on their semantic properties.

(213) a. Pedru laba. (Capeverdean) b. O Pedro lavou-se. (European Portuguese)

‘Peter washed.’
(214) a. Ana xatia. (Capeverdean)  
b. A Ana zangou-se. (European Portuguese)  
  ‘Ana got angry.’

(215) a. Porta abri. (Capeverdean)  
b. A porta abriu-se. (European Portuguese)  
  ‘The door opened.’

Contrast these with mata/matar/kill – in Capeverdean and English this needs a SELF-anaphor for the reflexive reading: Djon mata *(si kabesa); ‘Djon killed *(himself)’. In European Portuguese it is equally built with the se morpheme: O Djon matou-se.

4.2.1.2 Theta-roles and structural Case

In (213) the subject is assigned both the external θ-role, Agent, and the internal θ-role, Theme (these roles are assigned to different arguments in regular transitive entries).

In (214) only the internal θ-role, Experiencer, is assigned. Although we cannot have a passive construction in European Portuguese (*A Ana foi zangada), we have a semantic information about a cause or an external agent (even when it is not given in the sentence), and this becomes clear by the counterpart in English. This is not, thus, a case of inherent reflexivity, which would require an external argument, with a typical Agent θ-role.

In (215) only the internal θ-role, Theme, is assigned. By contrast with “real” unaccusatives, this type can alternate with transitive entries.

Notice that only the first, the entry in which the subject is assigned two theta-roles, has a reflexive reading.

The syntactic representations in Fiéis & Pratas 2004 are the following – (216) and (217) for European Portuguese and (218)-(221) for Capeverdean:
The DP *O Pedro* has two θ-roles, apparently violating the θ-Criterion: it gets the Theme θ-role in its base position (internal argument), and moves afterwards to Spec, VP (by copy and deletion). Then, in Spec,VP it will get/absorb the Agent θ-role. This constituent, now with two θ-roles, moves to Spec,IP where Nominative Case is assigned (contrarily to Reinhart’s proposal, there is no reduction in the valency of the verb, which remains a two place predicate). If we assume that Accusative Case is assigned later in the derivation, there might be no Accusative Case available here.

In European Portuguese, given the availability of a functional node for this, the *se*-type clitic is inserted in the morpho-phonological component of the grammar. This morpheme agrees\(^{162}\) with the subject (it must have the morphological shape of the subject person and number, although it reflects Accusative Case; this is a clue for viewing Accusative Case in European Portuguese as absorbed by this morpheme).

---

\(^{162}\) Maria Lobo (p.c. 2006) points out, crucially, that the best way to describe this relation is not by appealing to the co-reference between *se* and the subject (as we stated in the first version), but rather by saying that *se* agrees with the subject. Moreover, the functional label Refl seems superfluous, since the reflexive reading is indeed provided by the combination of two theta-roles. In (217), however, we have the *se* morpheme, and yet there is no reflexive reading (no combination of two theta-roles for the DP).
Like in (216), the se-type clitic in (217) agrees with the subject and is inserted post-syntactically. Contrarily to (216), in (217) the DP subject has only one θ-role, the internal one (Theme, in this case), which is the only one assigned by the verb, in a representation closer to a passive. There is no violation of the Case Filter: the internal argument is obligatorily moved to Spec,IP, and gets nominative case, just like in passives. In Fiéis & Pratas 2004 we also argued that under a compositional analysis this is not an unaccusative entry: semantically we have intuitions about the relation (established by the predicate) between the only argument and some Cause or Agent (external θ-roles) not expressed in the sentence. Their existence is however suggested by the Portuguese clitic se.

According to Kayne 1984 and Pesetsky 1995, this se clitic signals the presence of the external argument in the proper syntax. This is an argument in favor of an unaccusative analysis for these constructions. Our proposal differed from this one in a sense that the se morpheme is inserted post-syntactically. We considered that the only
valid argument for taking these constructions as unaccusative is the fact of them having a derived subject. Notice that comparing these with real unaccusative verbs to search for a possible contrast would not be productive, since the widely assumed unaccusative diagnostics do not apply to every verb in the same way. Moreover, the passivization test, for instance, reveals, if anything, that ‘The door opened’ is not an unaccusative entry: ‘The door was opened’ (passive of a transitive entry) / *’The tree was fallen’ (obvious ill-formedness of a passive of an unaccusative entry).
The reflexive reading in (218) depends on the contrast with the expected structure of a transitive verb like in (219). The absence of a DP – represented as a trace of the deleted DP (by copy and deletion, like in European Portuguese) – in the internal argument position gives way to the only interpretation possible: the one in which subject and object coincide. The DP subject has been moved to Spec,TP (the only functional projection in Capeverdean) carrying the complex θ-function Agent+Theme, and it gets Nominative Case. A se-type morpheme is not inserted, since there is no functional head available. We expect that Accusative Case has not been assigned given the absence of an overt internal argument. A compositional perspective is crucial here: there is only one overt argument that, together with the verb semantics, induces the reflexive reading.

In (220), the subject is generated in the internal argument position, receiving the internal θ-role – this can be Theme or Patient, depending on the semantics of the verb, as we could expect. The reading cannot be reflexive since the subject is not an Agent (it is widely assumed that subjects of reflexive entries are Agents). ‘The door’ does not ‘open itself’ (although we can say something like The door opened by itself). In European Portuguese, for a reading such as A Ana zangou-se a si própria, this anaphoric expression, of the SELF-type, would have to be generated in the internal argument position, with the subject Ana in Spec,VP.

In (221), the SELF-type anaphor si kabesa is obligatory: the reflexive reading for the verb mata ‘kill’ is at the extreme of improbability; whenever there is reflexivity it must be explicit in the sentence. Here this expression appears as an adjunct (and not the internal argument), assuring the reflexive reading, given that the absence of the internal argument of a transitive verb is not sufficient here. In European Portuguese, for this verb matar, the se morpheme is enough for the reflexive reading, without the anaphoric expression; se is inserted post-syntactically, on a functional head (which we have labeled as Refl, in order to account for the combination of two theta-roles of the DP).

As conclusions for this proposal in Fiéis & Pratas 2004, we have said that for the predicates under analysis the differences between Capeverdean and European Portuguese are not in the respective argument structure. The visible difference – the absence/presence of a morpheme of the se-type – depends on the architecture of the sentence in each of these languages. The relevant restriction is syntactic: the Split,IP
parameter. The functional structure of Capeverdean is restricted to TP, there is no place available for this type of functional categories. In this language, the reflexive reading is possible in certain contexts, when the internal argument of a transitive verb is missing. Also crucial is the assumption that \( \theta \)-roles are features and that movement to a \( \theta \)-position is allowed; dispensing with the \( \theta \)-criterion, it is possible to have a DP moving from one \( \theta \)-position to another, adding more than one \( \theta \)-role/feature on the same DP.

The labeling as unaccusatives does not seem relevant for the analysis of certain predicates, which do not behave according to some criteria traditionally used to test unaccusativity. Furthermore, this label does not hold for the same verbs cross-linguistically. A verb may have only an unaccusative entry in a language and be cumulatively transitive in the other: grow/crescer can be both unaccusative and transitive in English, whereas in European Portuguese or Italian it can only be unaccusative. (I grow flowers! * Eu cresço as flores).

We have defended that, for semantic reasons (that vary cross-linguistically as well), some particular entries of some transitive verbs like laba/lavar/wash allow the assignment of an internal \( \theta \)-role without assigning Accusative Case. In these environments, no other element – for instance the anaphoric expression of the SELF-type or some other information which allows the reflexive reading – is necessary for the sentence to be well-formed. With other verbs, some other element is obligatory. This is illustrated by the following contrast between sentences with djobe/ver ‘watch’, again in Capeverdean and European Portuguese:

(222)  
\begin{align*}
\text{a.} & \quad * \text{A Maria viu-se.} \\
\text{b.} & \quad \text{A Maria viu-se a si própria / ao espelho / na televisão.} \\
& \quad \text{‘Maria saw herself in the mirror / watched herself on tv.’} \\
\text{c.} & \quad * \text{Maria djobe.} \\
\text{d.} & \quad \text{Maria djobe si kabesa na tilibizon.} \\
& \quad \text{‘Maria watched herself on tv.’}
\end{align*}

This contrast shows that this European Portuguese clitic \textit{se} is not argumental. Otherwise, it would be hard to explain why (222a) is not accepted\textsuperscript{163}.

\textsuperscript{163} At the time we did not consider the following possible dependency: \textit{ver} ‘watch’ is a perceptive verb, thus non-agentive. True, in \textit{A porta abriu-se} ‘The door opened’ the subject is also not an Agent; but it is also an inanimate DP, which might allow for some different connections with non-agentivity. When we
In this subsection I have presented what is at stake with this empirical puzzle – reflexive verb entries that do not have an overt reflexive expression. I have also summarized briefly the lexicalist approach by Reinhart (there is the reduction of the internal argument of the verb) – which I had assumed in Pratas 2002 – and, with more detail, the joint work in Fiéis & Pratas 2004.

Further sources of Capeverdean evidence, however, have led me to extend the description of the dissimilarities between the reflexive sentences like (204b) and (206b), here repeated as (223) and (224):

(223) Maria laba.
     ‘Maria washed (herself).’

(224) Djon perde.
     ‘Djon got lost.’

The distinct English translations for the above sentences seem not to be an isolated phenomenon, but rather it is found among different entries. This distinction, however, does not have a direct correspondence to other specific semantic implications in their meaning. In the next subsection I will extend these previous works and check this possible organization into types of verbs.

4.2.2 A question of passivity

In this subsection I try to elaborate on a more semantically based analysis for the phenomenon under study, although at this stage this approach is not yet formalized.

This has departed from a tentative organization of the long list of predicates which allow for the lack of any morphological marker of reflexivity, in spite of having a reflexive reading. Notice that this does not include entries considered as unaccusative, such as Porta abri ‘The door opened’.

have human DPs as subjects of non-agentive verbs, some further needs arise; additional information is required in order to complete the meaning of the predicate. Take the following pair: *A Ana sentiu-se. vs. A Ana sentiu-se triste/ameaçada/observada. ‘Ana felt sad/in danger/observed.’

For an extensive discussion and analysis for the different Portuguese se morphemes (involving their classification as argumental and non-argumental, among other important details), see Duarte 2001.
This open list includes verbs that involve:

i) some change in one’s physical position: *deta* ‘lie down’, *xinta* ‘sit’, *labanta* ‘stand up’…

ii) some action/effect over one’s body: *laba* ‘wash’, *xuxa* ‘get dirty’, *modja* ‘get wet’, *pentia* ‘brush the hair’ / ‘get the hair brushed’, *fri* ‘get hurt’…

iii) some action/effect over one’s self (as long as it is different from *mata* ‘kill’): *perde* ‘get lost’, *kasa* ‘get married’, *ngana* ‘become mistaken’, *zanga* ‘become angry’…

iv) some type of self perception: *xinti* ‘feel’, *djobe na spedju* ‘see in the mirror’…

Whereas in the last set we have a type of states or activities, in the first three there is an associated telic property, for there is a focus on the result obtained. This resultative property is more visible in paraphrases in English where a *get* construction is used. Notice, however, that the verbs in the first set, involving some change in one’s physical position, are also translated into English without recourse to a *get* construction.

We can try to illustrate these effects in tables. Only Capeverdean and English verbs are included, since it has been through this specific comparison that I have felt the need of extending the semantically oriented approach. Therefore, the previous considerations regarding European Portuguese morpheme *se* have been left behind, for they are not conclusive (*se* is necessary in all these entries, obfuscating any possible insightful reflections about what lies beneath).

(225) i) entries that involve some change in one’s physical position

<table>
<thead>
<tr>
<th>Capeverdean</th>
<th>inherently reflexive</th>
<th><em>deta</em></th>
<th><em>xinta</em></th>
<th><em>labanta</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>SELF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>English</th>
<th>inherently reflexive</th>
<th>‘lie down’</th>
<th>‘sit’</th>
<th>‘stand up’</th>
</tr>
</thead>
<tbody>
<tr>
<td>get</td>
<td></td>
<td>'get up'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>become</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SELF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ii) entries that involve some action/effect over one’s body

<table>
<thead>
<tr>
<th>Capeverdean</th>
<th>inherently reflexive</th>
<th>laba</th>
<th>xuxa</th>
<th>modja</th>
<th>pentia</th>
<th>fri</th>
</tr>
</thead>
<tbody>
<tr>
<td>SELF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>English</th>
<th>inherently reflexive</th>
<th>‘wash’</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>get</td>
<td>‘get dirty’</td>
<td>‘get wet’</td>
<td>‘get hair brushed’</td>
<td>‘get hurt’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>become</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SELF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

iii) entries that involve some action/effect over one’s self

<table>
<thead>
<tr>
<th>Capeverdean</th>
<th>inherently reflexive</th>
<th>perde</th>
<th>kasa</th>
<th>ngana</th>
<th>xatia / zanga</th>
</tr>
</thead>
<tbody>
<tr>
<td>SELF</td>
<td></td>
<td>mata si kabesa</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>English</th>
<th>inherently reflexive</th>
<th>‘get lost’</th>
<th>‘get married’</th>
<th>‘get angry’</th>
</tr>
</thead>
<tbody>
<tr>
<td>get</td>
<td></td>
<td>‘become mistaken’</td>
<td>‘become angry’</td>
<td></td>
</tr>
<tr>
<td>become</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SELF</td>
<td></td>
<td></td>
<td>‘kill oneself’</td>
<td></td>
</tr>
</tbody>
</table>

iv) entries that involve some type of self perception

<table>
<thead>
<tr>
<th>Capeverdean</th>
<th>inherently reflexive</th>
<th>xinti</th>
<th>djobe na spedju</th>
</tr>
</thead>
<tbody>
<tr>
<td>SELF</td>
<td></td>
<td></td>
<td>djobe si kabesa na tilibison</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>English</th>
<th>inherently reflexive</th>
<th>‘feel’</th>
<th>‘look at the mirror’</th>
</tr>
</thead>
<tbody>
<tr>
<td>get</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>become</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SELF</td>
<td></td>
<td></td>
<td>‘watch oneself on tv’</td>
</tr>
</tbody>
</table>
In these tables we easily observe that within the types ii) and iii) we get the most ‘get’ and ‘become’ constructions in English as corresponding to Capeverdean inherently reflexive entries (except for the verb ‘wash’, in type ii) – this verb is in fact one of the most mentioned in studies about these reflexive entries, such as in Reinhart 1996).

This observation turned out to be of crucial influence to my last view on the puzzle. In trying to distinguish the theta-roles assigned to the subject DPs, we may have an Agent in the case of laba ‘wash’ and an Experiencer in the case of xinti ‘feel’. However, if we consider the paraphrase for laba as ‘get washed’, we obtain a Theme-like theta-role. This applies whatever the analysis, but we shall see the origin of the DP (where it is base-generated) might be an important question here.

It has hitherto been rendered clearer that in these reflexive-like constructions with this type of verbs the subject has a Theme/Patient theta-role. Syntactically, this means it has been generated in the internal argument position. This has been captured in (218), as proposed in Fiéis & Pratas 2005.

Curiously, when we take the verb xuxa, which could be semantically related to laba (both in group ii)), we arrive to a core point in the whole discussion: the internal theta-role of the DP when an accident happened is not the same as when there is a agentive reading, and the latter must not be the same as when someone else specifically xuxa or laba the person in question (in which case we would have two distinct DPs, in a transitive configuration). Either way, there is still something in all these hypotheses that is focused on the result.

This implies two different facts:

i) there is a crucial, although subtle, distinction between the internal theta-roles of these constructions and the internal theta-role in transitives/regular passives;

ii) there is some telicity involved – we are referring to the accomplishment and not to the process or activity.
This last fact could apparently be tested with the change from the telic form in PastPer to the atelic in any form of progressive. However, we observe that this change in grammatical aspect has no implications in the relevance of the event result\(^\text{164}\).

Consider the following sentence:

(226) *Kantu-N txiga kasa, Pedru sata lababa.*

‘When I arrived home Pedru was getting washed /?? washing.’

The progressive reading is evident, and yet, as we observe by the preferred translation in English, the sentence is focused on the result.

This Capeverdean reflexive reading in a progressive environment even resists sluicing. It does not hold for the English versions which are not focused on the result:

(227) a. *Algen sata lababa mas N ka sabe e kenha.*

?? ‘Somebody was washing but I don’t know who.’

b. *Algen sata pentiaba ku nha skova, mas N ka sabe e kenha.*

‘Somebody was ?? brushing with my hairbrush but I don’t know who.’

In Capeverdean, under VP ellipsis (antecedent-contained deletion), we get a sloppy reading (as for English, it must be investigated if this holds for different tenses):

(228) *Pedru laba / ta laba tudu dia y Maria tambê.*

‘Pedru washed/washes everyday, and Maria does too(washes Maria/*washes Pedru).’

In Capeverdean, the reflexive reading is likewise maintained in embedded non-finite environments, such as in causative constructions:

\(^{164}\) The fact that the verbalized temporal construction seems to be referring to a process does not mean that we are not focused on the result that will come out of it. In European Portuguese, for instance, it is possible to apply a progressive construction to a typically telic predicate, like chegar ‘arrive’ (an achievement).

(i) *Estou a chegar.*

‘I am arriving.’

This means that I am close to the achievement point – to which the telic property is associated – and not that I am equally arriving in successive moments in a continuous period of time.
When we try to isolate the environments where the DP clearly has an active involvement in the event from the ones where it is better understood as a Theme, the Capeverdean construction stays the same:

(230)  

a. Onti txobe txeu y limarias modja.  
     ‘Yesterday it rained a lot and the animals (in the backyard) became wet.’

b. Katxor modja p’-e fika mas fresku.  
     ‘The dog got wet so that he becomes refreshed.’

The two English translations above suggest a further interesting fact: there is a distinction between the environment where the subject is supposed to be innocent (a become-construction) and the environment where the subject must have had some active intervention (a get-construction).

With human DPs, when there is an active involvement in the action (such as in (232)) we may say clearly the subject has an Agent theta-role.

(231)  

Pedru sta la na rua ta brinka ku agu. N ta pusta m’-e sata modja.’  
     ‘Pedru is outside playing with water. I bet that he is becoming wet.’

(232)  

Pedru sata modja (di xokota) p’-e fika mas fresku.  
     ‘Pedru is getting wet (on purpose) so that he becomes refreshed.’

As we see by the Capeverdean contrast between the matrix and the embedded clauses, in (232), there is a different form for saying that he is getting refreshed – this goal is presented with fika, literally ‘stay’. This occurs when the clause at stake refers exclusively to the result and not to any action leading to it; this is the case with the clause p’-e fika mas fresku, since the action leading to this result is contained in a different clause, the matrix clause. The relevant translation into English shows that there is an agentive-type direct implication of the person in the verb ‘get’, whereas the implication in the verb ‘become’, if any, is more indirect. Even in ‘Pedru got lost (in
the city), there is some presupposition that Pedru did something which led to him getting lost.

Thus, in some of these predicate constructions, although we may not be referring to any motivation behind the action but rather to the result itself, there is the presupposition (not always an information clearly stated – and this is the crucial difference from a true reflexive) that the theme affected by the action had some active responsibility in the result as well.

Thus, we have two possible environments, which assume the same form:
- either the subject has been /is being submitted (or whatever the tense and aspect in question) to some event which results in a change of state for it / him;
- or it / he has performed / is performing (or whatever the tense and aspect in question) something which results in a self change of state.

Consider some more examples from the group (ii) – the ones which perform some effect over one’s body:

(233) Maria, bu sata xuxa, bu ka ta ba festa.
     ‘Maria, you are getting dirty, [like this] you can’t go to the party.’

(234) Bu sata pentia otu bes? Kantu bes ki bu pentia oxi?
     ‘Are you brushing your hair again? How many times did you do it today?’

(235) N sata muda kel movel-li y N sata magua.
     ‘I am moving this piece of furniture and I am getting hurt.’

(236) Kenha ki ta laba tudu dia?
     ‘Who washes [reflexive] everyday?’

Now consider some examples with verbs from group (iii) – the ones which refer to some effect over oneself. The verbs in question are perde ‘get lost’ and xatia ‘become upset’.

236
(237) Kenha ki perde onti?
    ‘Who got lost yesterday?’

(238) Di purmeru bes ki e ben Lisboa, e perde.
    ‘In the first time he came to Lisbon, he got lost.’

(239) Ka bu xatia, Madalena!
    ‘Don’t you get upset, Madalena!’

One possible approach to these might be by way of seeing them as middle constructions, or as having the meaning (although not the structure) of a passive identical to, for instance, the German “process passives” – “Vorgangs” (see Kratzer 2000) –, which are built with the auxiliary werden (‘get’, ‘become’), and not with sein ‘be’ (these ones, with ‘be’, are the adjectival passives). We have seen, however, that between a ‘get’ effect and a ‘become’ effect there are some distinctions as well.

Considering that, from groups ii) and iii), only a few English verbs, such as ‘wash’, allow for these constructions with a similar reading (and not in all circumstances, as we have just seen – see (227)), we might also say that the set of verbs that allow these constructions in English is a subset of the verbs that allow them in Capeverdean (as said before, this empirical phenomenon is much more productive in the latter).

The question left open for now is on whether all the verbs in the type-listing, above – and not only the ones in ii) and iii) –, can be considered as some sort of “process passives”: either with the ‘become’ reading (Limarias modja na txuba) or with the ‘get’ reading (Pedru modja p’e fika mas fresku).165

The clear exception is hitherto djobe na spedju ‘see oneself in the mirror’, which implies a specific context: Pedru got seen in the mirror triggers a true passive reading (to which a by-phrase could be added). My explanation for this very specific context (recall that for ‘watch oneself on tv’, for instance, Capeverdean requires the SELF-

165 Under a feature-driven approach to movement, one of the more relevant consequences of this proposal for the structure of this type of constructions is that a passive does not project vP; if we take this functional projection as being responsible for Accusative Case, we know now that in these environments there is no Accusative Case feature at all. The object has to move to Spec,TP and check Nominative Case. I will also leave this question open.
anaphor: *Maria djobe si kabesa na tilibison* the “reflexive” reading involved is provided by pragmatic factors: what we know about the world allows us to know that when someone looks at a mirror the image that is seen is, by default, the seer’s own image.

Now let me briefly present some contrasts involving animated and non-animated DPs.

(240)  

a. *Limarias modja na txuba.*

‘The animals became wet with the rain’

b. *Porta modja na txuba.*

‘The door became wet in the rain.’

c. *Porta abri.*

‘The door opened.’

We observe that a predicate like *modja* does not preclude, in Capeverdean, the same type of construction for a non-animated DP (it might be expected that it required a true passive). In English, again, we are able to observe the contrast between this *become*-construction in (240b) and an unaccusative-like construction in (240c):\(^{166}\)

This type of “process passives” is, thus, not sensitive to animacy, which renders it clear that the DP did not have, in this case, any agentive contribution. This is more visible when we try to apply a different paraphrase in English: *‘The door got wet.’*

Recall that in European Portuguese all the entries corresponding to these must have a *se* morpheme. In unaccusative-like constructions, such as ‘The door opened’ and ‘The glass broke’, the *se* morpheme is sometimes optional (recall that in Kayne 1984 and Pesetsky 1995 the morpheme *se* was viewed as evidence of the external argument in the syntax).

For a future line of inquiry, it must be taken into consideration the fact that we have, lying underneath the visible argument structure of these reflexive predicates – which also have transitive entries –, a complex matter of semantic interactions which involve the following elements:

\(^{166}\) These constructions are referred to as “unaccusative-like”, since, as said earlier, these verbs do not pass all tests for unaccusativity. One of them is, obviously, the impossible passivization. Check the following contrast: *‘The tree was fallen’, but ‘The door was opened’.*
a) animated / non-animated DPs;
b) presupposition of some action of the DP for the event result to obtain (either on purpose, which corresponds more to a *get*-reading, or accidentally/involuntarily, which corresponds more to a *become*-reading);
c) this result falling on the DP itself (with the further constraint of the verbs in group iii), like *perde*, *zanga* or *ngana*, being slightly more restricted to humans).

A distinct example of associations between certain predicate Aktionsart properties and the licensing of a certain type of passive is given in Borer 1998 on impersonal passives, where the external argument is eliminated or reassigned to a by-phrase, but no direct object moves to subject position. These are generally intransitive (unergative or unaccusative) verbs, or intransitive versions of verbs (such as ‘dance’, ‘swim’, ‘drown’) that in certain contexts may assign Case to an object (like transitives).

In English there are no impersonal passives of this type (although there are impersonal passives with complement clauses, just like ‘It was said that…’), and it has been argued that this is due to the following association: if a verb does not assign Accusative Case (be it unergative or unaccusative) it may not receive passive morphology. Confront, on one side, English: *‘It was swum here frequently’* and *‘It was drowned here frequently’* with, on the other side, the case of Dutch, which allows impersonal passives with unergatives, but not with unaccusatives: *In de zomer wordt er hier vaak gezwommen* (in the summer it is here frequently swum), *In de zomer wordt er hier vaak verdronken* (in the summer it is here frequently drowned).

In the next subsection I will address the rules and constraints involved in the word order inside ditransitive constructions, more specifically the order between the two complement DPs. What will be presented is, as in all sections in this chapter, a summary of a joint work, this time Fiéis & Pratas 2005. In the last subsection, some new arrangements will be suggested.
4.3 DOCs versus to-datives

This section is about Double Object Constructions, the so-called DOCs. In subsection 4.3.1 a Capeverdean problem will be presented. Subsection 4.3.2 summarizes some previous proposals on whether or not there is an underlying relation between DOCs and dative constructions with preposition (to-datives). Subsection 4.3.3 presents what has been proposed in Fiéis & Pratas 2005 in order to account for some specific Capeverdean puzzles.

Finally, in subsection 4.3.4 a different approach is proposed, always in the view that some empirical puzzles are left open for future research. This, however, does not interfere with the main goal of this section: to illustrate some complexities inside Capeverdean verb domain, most of them determined – as expected – by the idiosyncrasies of the predicates.

4.3.1 Capeverdean puzzle

In Capeverdean there seems to be no transformational relation between DOCs (Double Object Constructions) and to-datives, as some authors have proposed for the English pair in (241):

(241)  a. John gave Mary the map
       b. John gave the map to Mary.

When a PP is involved in a ditransitive construction in Capeverdean, a different meaning obtains: some purposive content becomes involved. This fact, which seems due to the specific semantic contribution of the preposition pa (the only one available in the language in order to test the contrast between DOCs and some kind of to-dative counterpart) leaves many questions open. Namely what is the true value of to in English, like in (242) and whether this value is independent of the one in to-datives167:

(242) I left early to avoid the traffic.

---

167 Some English words like *to* or *for*, as also *pa* in Capeverdean, seem to have some different lives (prepositions, complementizers, etc.), according to the environment they occur in; sometimes, however, it is not that easy to distinguish which properties are at stake in a given environment of each of these little words.
Before proceeding to the possible problems, let me present the relevant Capeverdean data. As stated in Baptista 2002, Capeverdean DOCs always display the order Verb-Indirect Object-Direct Object (henceforth V-IO-DO).

This applies to structures V-DP-DP, as in (243), as well as to V-DP-Pron, V-Pron-DP and V-Pron-Pron, like in examples from (244) to (246):

(243)  a. *N da Manel dinheru
       I   give Manel money
       ‘I gave Manel the money’
   b. *N da dinheru Manel

(244)  a. N da Manel el
       I   give Manel 3SG
       ‘I gave it to Manel’
   b. *N da-l Manel

(245)  a. N da-l dinheru
       I   give-3SG money
       ‘I gave him the money’
   b. *N da dinheru el

(246)  a. N da-l es
       I   give-3SG 3PL
       ‘I gave them to him’
   b. *N da-s el\(^\text{168}\)

In the DP-PP clauses, besides being a lot less productive with a predicate like \textit{give}, a Purpose must be involved, as we can see in the contrasts in (247) and also in (248):

(247)  a. N da Maria livru.
       I give Maria livru
       ‘I gave Maria a book.’

\(^{168}\) This sentence is well-formed with the meaning ‘I gave it to them’. 

b. *N da livru pa Maria*

I give book for Maria

‘I gave the book to Mary [for Mary to have it] / [for Mary to read]’

(248) a. *N mostra Pedru konta.*

I show Pedru check

‘I showed Pedru the (restaurant) check.’

b. *N mostra konta pa Pedru.*

‘I showed the check to Pedru [for him to pay it].’

### 4.3.2 Previous proposals

The following are some of the previous proposals centered in English data.

For Kayne 1984, both arguments in DOCs are embedded in a small clause, as in the structure in (249).

(249) \[
\begin{array}{c}
\text{VP} \\
\text{V} \quad \text{SC} \\
\quad \text{give} \quad \text{NP} \quad \text{NP} \\
\quad \text{Mary} \quad \text{a book}
\end{array}
\]

In the ditransitive constructions with preposition only the first NP behaves as the argument of the verb, resulting in a structure like (250):

(250) \[
\begin{array}{c}
\text{VP} \\
\text{Subject} \quad \text{V'} \\
\text{V'} \quad \text{PP} \\
\text{V} \quad \text{NP} \quad \text{to Mary} \\
\quad \text{give} \quad \text{the book}
\end{array}
\]
A different approach, in Larson 1988, posits a hierarchical structure for to-datives involving VP-shells, in which the NP direct object is in the specifier position of the lower VP. The verb raises to the higher V for Case and agreement reasons.

(251) \[
\begin{array}{c}
\text{VP} \\
\text{V'} \\
\text{V} & \text{VP} \\
\text{send}_i & \text{NP} & \text{V'} \\
\text{a letter} & \text{V} & \text{PP} \\
& \text{t}_i & \text{to Mary} \\
\end{array}
\]

(Larson 1988: 343)

The author also states that DOCs are derived from to-datives. “A derivational approach to the dative-double object relation is clearly desirable under any strong theses about the relation between structure and assignment of thematic roles.” (Larson 1988: 350)

In the DOC, as in (252), contrary to the ditransitive construction with preposition, the DP a letter is base generated in adjunction to the lower V’. In order to derive the surface order, the verb raises to the higher V position and the NP Mary raises to the specifier position of the lower VP.

(252) \[
\begin{array}{c}
\text{VP} \\
\text{V'} \\
\text{V} & \text{VP} \\
\text{send}_i & \text{NP} & \text{V'} \\
\text{Mary}_j & \text{V'} & \text{NP} \\
\text{V} & \text{NP} & \text{a letter} \\
& \text{t}_i & \text{t}_j \\
\end{array}
\]

(Larson 1988: 353)
As said earlier, Larson’s analysis considers both constructions as structurally related: the ditransitive with preposition is the underlying order (in which the UTAH is respected), and the DOC is derived through a strictly syntactic operation\(^{169}\).

Larson 1988 is not the only one defending a monossemic approach, motivated more for syntactic than for semantic reasons. For other monossemic proposals see Baker 1988, Dowty 1978, a.o. (these authors consider, like Larson 1988, that the variant with preposition is the basic one and the double object is derived from that one), and also Dryer 1986 and Kiparsky 1985 (these authors consider that the basic order is the double object one). There are other authors (like Butt, Dalrymple & Frank 1997) that propose a non-derivational perspective: the same meaning generates two different argument expressions.

On the “opposite” side, there are the polissemic approaches, according to which the two variants result from different verb meanings. Some examples of these are: Hale & Keyser 1996, Harley 1997, Krifka 1999, 2001 and Speas 1990, a.o.. These proposals generically point out a mapping of the DO in each variant according to the different semantic conditions of the verb. Hence, we have (adapted from the description in Levin & Rappaport Hovav 2005\(^{170}\)):

\[
\begin{align*}
i) & \quad \text{‘x causes y to have z’, generating the order DOC (241a)} \\
ii) & \quad \text{‘x causes z to be in y’, generating the variant with preposition (241b)}
\end{align*}
\]

These relevant examples are repeated in (253):

(253) a. John gave Mary the map  
       b. John gave the map to Mary.

Hence, the affected or “caused” argument is y in the first case and z in the second. In both, the two NPs – y and z – are internal arguments (none appears in adjunction).

---

\(^{169}\) The DOC is base generated in conformity with the UTAH – Uniformity of Theta-role Assignment Hypothesis – and transformed via movement in a structure that is not in conformity with the UTAH. For a discussion, under a different view, of DOCs in creole languages, see Martinez-Sanz 2006.

\(^{170}\) Note that these authors defend the monossemic approach to English; this polissemic description is only a side step in order to defend their position.
Finally, for Pesetsky 1995 the two structures (give Mary a book and give a book to Mary) are not derived from one another. We can see this in (254) and (255).

(254) Double complement structure (adapted from Pesetsky 1995:174)

(255) DOC (adapted from Pesetsky 1995:197)

Pesetsky 1995 shows that the differences in the interpretation of the two sentences result from differences in the semantics of the two prepositions. In the DOC the Goal θ-role is directly selected, whereas in the double complement construction (the to-dative) the
selection of the Goal is indirect, and depends on the properties of the prepositions and not on the construction itself.

As for those cases in which the θ-role is assigned to a non-animated DP, as it occurs with the verb *send*, forbidding the DOC in English, we must retain the idea, also expressed in Pesetsky 1995, that the different acceptability of the sentences in (256) can be explained by the fact that “the semantics of to-objects seem to be a superset of the semantics of directly selected Goals” (Pesetsky 1995:141).

(256) a. *I sent a letter to Lisbon.
    b. *I sent Lisbon a letter.

4.3.3 Our proposal

Notice that in Capeverdean, notwithstanding the question of there being a different meaning involved in the *to*-dative \(^{172}\), there is some other empirical fact that must be accounted for: in certain constructions with ditransitives the IO is obligatory while the DO is not.

Let us look at the example in (257), where a discourse bound DO is absent from the embedded clause. Notice that both the IO and the DO might be easily recovered from the context; hence it is not clear why only the IO is obligatory.

(257) *Si nha fidju-femia kreba odja fotos, N ta mostraba *(el) / *es \(^{173}\)
     ‘If my daughter wanted to see the photos, I would show her.’

As for the syntactic structure of ditransitive constructions in Capeverdean, also an evident fact is that the IO c-commands the DO, thus explaining the contrast in (258).

---

\(^{171}\) For further development on this issue, see Harley 2002, where G is replaced by P\(_{\text{HAVE}}\).

\(^{172}\) As stated above, these *to*-datives with verbs like *da* ‘give’ are not much productive in Capeverdean, being mainly produced by my informants in elicitation interviews.

\(^{173}\) The ill-formedness with *es* seems due to some odd morpho-phonologic effects of the simultaneous presence of both pronouns being strongly avoided, and since the first one is obligatory, this second one (the DO) is undesirable.
(258)  a. N da kada mininu si livru.
   ‘I gave each boy his book.
   b. *N da kada livru si autor.
   ‘I gave each book his author.’
   c. *N da si autor kada livru.
   ‘I gave each book his author.’

In (258a) everything is okay: the order is IO-DO and the bound possessive is c-commanded by its antecedent. In (258b), where the possessive could also be c-commanded by its antecedent, there is an illegitimate DO-IO order. In (258c) we see more clearly that IO c-commands DO; the bound expression in the IO position violates principle A.

Given all this, in Fiéis & Pratas 2005 the proposal for the configuration of Capeverdean DOCs is as follows:

i) VP is embedded in vP;
ii) IO is in Spec, VP;
iii) DO is in the complement position.

We also considered that in Capeverdean the Spec,VP position must be filled\(^{174}\), meaning that no object generated there can be null. This would explain the impossible null DO in transitive constructions as in the answer in (259), below – check the contrast with the embedded DOC in (257), where the obligatory DP is the IO.

In (259) the DO is the only internal argument of the sentence. This means, we proposed then, that in transitive constructions the DO is in this crucial Spec,VP position – N da-l – and the DP that receives the Goal \(\theta\)-role (a possible \(\theta\)-role for the IO in a DOC) is here inside an adjunct PP – pa nha mai:

(259) Q. Dja bu da Manel dinheru?
   ‘Have you given Manel the money?’

\(^{174}\) This could extend to some language specific prohibition of empty specifiers, supposing the specifier of T cannot be empty either (there are no referential null subjects in the language). However, my current proposal for the specifier of T is that when it is not needed, it is not projected; maintaining a coherent approach for different specifiers would not be coherent with the positing of an obligatory DP on Spec,VP; in other words, why would it be that also Spec,VP is always projected, obligating the DP presence?

‘No. I gave it to my mother.’ / ‘I gave to my mother.’

We proposed for Capeverdean DOCs the underlying structure in (260):

(260) Capeverdean DOC: N da Maria livru. ‘I gave Maria a book.’

```
  vP
     |
   N   v'
     |
     v  VP
        |
     da_i NP  V'
        |
       Maria V  NP
        |
           |  livru
```

For the constructions with *pa* (recall that we assume they are not semantically equivalent to DOCs), we proposed the following underlying structure:

(261) Capeverdean to-dative: N da-l /N da dinheru pa nha mai. ‘I gave it to my mother.’

```
  vP
     |
   N   v'
     |
     v  VP
        |
     da_i -l/dinheru_j V'
        |
       V'  PP
          |
         V  NP  pa nha mai
          |
             |  |  |
             t_i  t_j
```
We left for further investigation the idiosyncrasies of predicates with respect to the Goals they admit and the positions they can assume; recall that in English it is impossible to have an inanimate object as a directly selected Goal with the verb ‘send’, as in *I sent Lisbon a letter.

The question is that the proposal on the dissociation of the DOC semantics from the one in the to-dative “counterparts” holds only for certain predicates (like da ‘give’, for which in the DOC no Purpose seems involved whereas the to-dative adds this meaning). Consider the following pair, where this distinction seems not to hold:

\[(262) \quad \begin{align*}
    a. & \quad N \text{ manda } nh\text{'armun un prenda}. \\
    & \quad \text{‘I sent my brother a gift.’}
    \\
    b. & \quad N \text{ manda un prenda pa nh\text{'armun}.} \\
    & \quad \text{‘I sent a gift to my brother.’}
\end{align*}\]

We might say that manda ‘send’ involves a Purpose in the DOC already (this is, as expected, determined by the semantics of the predicate), hence there is no such “adding” in the to-dative.

### 4.3.4 Some new arrangements

The following are the major lines for future inquiry on this matter:

- i) which ditransitive verbs allow for these or those relations?
- ii) can they be organized according to an established typology (like the ones which imply Purpose and the ones which do not)?
- iii) how can these distinctions be captured in any structural configuration?

Nevertheless, in the light of what I am defending in this dissertation on Capeverdean architecture, some new arrangements must be added to a possible underlying structure of the DOCs and the ditransitives with preposition of the verbs studied, like da ‘give’.

One of the questions left unanswered at the time is twofold:

- i) what happens to Case for both DPs (Indirect Object and Direct Object)?
- ii) why does the Direct Object move to the specifier of V in (261)?
In other words, with respect to ii): having the DP in the specifier of V in this construction is very useful in order to account for its impossible deletion; we might assume, as said before, that this position cannot be empty in the language.

Assuming, however, that when a specifier is not needed it is not projected, it is not clear why the DP moves there; the reason for this movement (hence, the reason for this specifier to be needed and, thus, for it to be projected in the first place) must be in the DP. The only possible reason could be related to Case.

Moreover, recall that we must explain why, in a Double Object Construction, the Indirect Object is obligatory and the Direct Object is not, since both of them are easily recovered from the context, as in (257), here repeated as (263):

(263) Si nha fidju-femia kreba odja fotos, N ta mostraba *(el) / (ex)
    ‘If my daughter wanted to see the photos, I would show her’

Given this crossed arguments and empirical puzzles, and maintaining that the constructions in question are not semantically equivalent, I will essay a simplified version of the proposal in Larson 1988 for DOCs (check (252)) and will be back to the proposal in Kayne 1984 for the pa-construction\(^{175}\) in (250).

Starting with the DOC, we have an adjunction configuration. This means the following:

i) no DP complement is in Spec,VP;

ii) the IO is in the complement position of the lower V, getting object case;

iii) the DO (theme) is in adjunction to the lower V – also in a sisterhood relation that allows for the morphological object case to be assigned;

iv) IO and DO are distinguished in terms of theta-roles;

v) I am referring to case here not in terms of abstract features, but rather as a morphological marking (associated to its grammatical function); hence, the same verb has determined in its own semantics the possibility of selecting for two objects, given that they are in the necessary syntactic position (recall that morphology is fed by syntax) – thus respecting UTAH.

\(^{175}\) Calling to-dative to this would not be correct, since in Capeverdean there is no morphological distinct marking for Accusative and Dative, only for pronominal subjects and objects.
vi) I will leave open the question about abstract case.

The adjunction configuration might help explain why in DOCs the DO is dispensable when it is perfectly identified by some discourse information.

The sentence is taken with pronominal forms, so that morphological object case is visible (recall that in the language there is no such marking for full DPs); for reasons of clarity, suppose that it stands for *N da nhas primu livru* ‘I gave my cousins a book’.

(264) Capeverdean DOC: *N da-s el:*

```
VP
  /
/   \
N    V'
  /
/   \
V'    DP
  /
/   \
V    DP
  /
/   \
da -s  el
```

One of the problems regarding this could come from the facts in (258), since a c-commanding relation between both objects no longer holds; hence, it would not be clear how *N da kada mininu si livru* is permitted (there are no Principle A effects).

We must admit, however, that the binding of a DP in an adjunct must be possible as long as the antecedent precedes the bound DP, since we have the following contrast:

(265) a. *Kantu e_j^*_i era pikinoti, Pedru_j ta koreba txeu.*
    b. *Kantu Pedru_j era pikinoti, e_j^*_i ta koreba txeu.*
    c. *E_j^*_i ta koreba txeu kantu Pedru_j era pikinoti.*
    d. *Pedru_j ta koreba txeu kantu e_j^*_i era pikinoti.*

‘When Pedru was a little boy, he used to run a lot.’

In (265), no matter the order relation between matrix clause and the adjunct clause, the co-reference is only possible when the R-expression precedes the pronoun, in (265b) and (265d), thus avoiding Principle C effects (this is confirmed with object DP/pronoun:
Kelora ki N odja-l Pedru staba na jardin. / Kelora ki N odja Pedru, eij staba na jardin. ‘I the time I saw him, Pedru was in the garden’). Therefore, what really is at stake is the order between the antecedent and the bound pronominal.

As for the pa-construction, in the light of this there seems to be no reason not to accept a structure like the one proposed in Kayne 1984, as illustrated in (250), here repeated (and adapted for Capeverdean) as (266):

(266) Capeverdean ditransitive with preposition: N da-l pa bo.

The difference here regarding object morphological case is that the second DP gets it from the preposition, and not under a sisterhood relation with V.

Basically, thus, this is a transitive structure to which a PP has been adjoined. In order to show that the adjunction configuration does not preclude a binding relation between both DPs, we have the following contrast (in (267b), Principle A effects arise):

(267) a. N da kada livru, pa si autor.
   b. N da si livru pa kada autor.

This section has been about Double Object Constructions in Capeverdean. In 4.3.1, I have described the Capeverdean puzzle regarding ditransitive constructions – the ditransitive with preposition is nor semantically equivalent to the Double Object Construction (preposition pa seems to involve some sort of Purpose). In 4.3.2 I have summarized the previous proposals for the pairs DOC – to-datives – that have been
considered in Fiéis & Pratas 2005, and in 4.3.3 I have summarized our own proposal at the time. Finally, in 4.3.4 I have presented some new arrangements in the light of the present dissertation: both constructions are accounted for by an adjoined configuration, in the first one the adjoined node is a DP, in the second one the adjoined projection is a PP; in both, the object DPs can be assigned morphological case, as well as the respective theta-roles.

4.4 Conclusion

In this chapter I have addressed three different puzzles on internal arguments, trying to account for them in a perspective coherent with the concerns that have been guiding the present dissertation, in particular the ones concerning economy and simplicity.

The realization of object pronominals in different forms depending on the presence of postverbal \(-ba\) has been addressed in 4.1: the proposed solution has been based on the one previously essayed in Costa & Pratas 2003, further developed in Pratas 2003b and completed in Pratas & Salanova 2005; it contends that the prohibition of the object enclitic in the presence of the temporal affix \(-ba\) must be accounted for by way of a morpho-phonological approach.

The absence of an anaphoric expression in certain reflexive contexts has been addressed in 4.2, departing from the proposals considered in my Master’s dissertation and using the developments achieved in Fiéis & Pratas 2004; the main lines in this last joint work included dispensing with the theta-criterion (in the fashion proposed in Horsntein 1999) and assuming that in certain reflexive entries the DP subject is generated as internal argument and then moves to the subject position, combining two theta-roles; the developments that I present here go further in the organization of such verbal entries and discuss, under a more semantically oriented perspective, their similarities with a type of passives that in English are marked by \textit{get} and \textit{become}. In 4.3, I have developed the joint work Fiéis & Pratas 2005, on the structure of ditransitives.

In the next chapter I will investigate null subjects in finite environments in Capeverdean (the null category that has been assumed as the subject of non-finite clauses will be addressed in Chapter Six). Both null referential subjects and null expletive subjects will be considered, as well as the so-called Null Subject Parameter.
Chapter Five

Null subjects in finite clauses

Subjects have always been a special topic in generative research, particularly for the intuition (mainly grounded on English and French data) that sentences could not exist without them. Nevertheless, consider the following contrast:

(268) a. *(It) is raining.  
   b. *(Il) pleut.  (French)  
   c. Está a chover.  (European Portuguese)  
   d. Sata txobe.  (Capeverdean)

Whenever the subject position is not filled by a DP with a theta-role, English and French require “something” to be inserted in the subject position, even though this “something” brings no semantic contribution to it. This contrasts with romance languages such as European Portuguese (as well as Spanish and Italian), which allow a silent subject position in these expletive constructions (the implications that come with this “silence” are a matter of great debate).

Furthermore, European Portuguese (as well as Spanish, Italian and other languages) allows for a true DP subject to remain silent, whereas in English (and French, and others) this is not possible.

(269) a. Comeram o bolo.  (European Portuguese)  
   b. * Ate the cake.  
   c. * Kome bolu.  (Capeverdean)

Under the Principles and Parameters framework, this contrast between English and French-type languages, on one side, and the Italian and European Portuguese-type, on
the other, has been useful to feed the presentation of what could be called the prototypical parameter: the pro-drop parameter. Languages in which the subject pronoun can be dropped (or null) are pro-drop languages (they have a positive value for this parameter), the ones that require a subject pronoun to be heard are not (they have a negative value for the parameter).

The ultimate goal of this chapter is to present striking arguments in favor of the separation between null referential subjects (the null category pro) (269) and the so-called null expletive subjects (expletive pro) (268).

With respect to the latter I intend to demonstrate that, at least in Capeverdean, they do not exist. In other words, in expletive constructions, where no expletive expression is inserted in the subject position (such as it is in English or French), nothing needs be in Spec,TP, and, thus, this position is not projected. Recall from earlier chapters that, following minimalist concerns, the major structural principle guiding this dissertation is that only the necessary functional positions are projected. As for null referential subjects (pro), the goal is to provide empirical evidence showing that Capeverdean is not pro-drop: it allows null referential subjects in finite clauses only in certain bound embedded contexts (Costa & Pratas 2006).

In section 5.1, the relevant Capeverdean contexts with respect to the null referential subject hypothesis will be presented, both in matrix (5.1.1) and embedded clauses (5.1.2).

In section 5.2 I will sum up the theoretical evolution on the topic of the null subject parameter, both on the proposals for what is involved in a positive value for it, like the licensing role of rich verbal agreement morphology, and the possible relation to other types of parameters. In section 5.3 the expletive contexts will be addressed: crucial data will be provided (5.3.1) and the EPP will be discussed (5.3.2). In 5.4 I shall summarize the lines of reasoning that support the consideration of Capeverdean as a non-pro-drop language, as well as the relevant conclusions of this chapter.
5.1 Capeverdean: it depends

Whereas expletive constructions with no overt expression in subject position constitute a widely attested phenomenon in Capeverdean, null referential subjects are categorically ruled out in most Capeverdean sentences, even in embedded finite clauses where their semantic content could be easily recovered from the context.

Crucially, null referential subjects are disallowed even in some non-finite embedded clauses; this phenomenon shall be addressed as a separate topic in the next chapter.

Only in specific embedded clauses, where the empty subject position is a variable bound by an operator, can we find clear null referential subjects. This much specific environment, as we shall see, is not sufficient to posit Capeverdean as pro-drop.

The impossibility of null referential subjects in matrix clauses is supported by the data in 5.1.1. My first proposal (Pratas 2002) has been to consider Capeverdean as a semi-pro-drop language: it arguably allows all types of null expletive subjects and no null referential subjects. Nevertheless, both the empirical results obtained in fieldwork and the theoretical approaches that I have contacted with since then have led to a more complex view with respect to that “semi-pro-drop” observations.

The differences between my previous view – and mainly the “semi-pro-drop” classification – and my current perspective are twofold:

i) to have no overt expletives has no consequences on the pro-drop status, since these are phenomena of a different nature (although English data – no null referential subjects in finite contexts, obligatory overt expletives – may have had a bewildering effect over the years);

ii) Capeverdean allows, actually, for a few null referential subjects. In Costa & Pratas 2006 these particular embedded contexts have been discussed. The relevant data and analysis are in 5.1.2.

5.1.1 Matrix clauses

In this subsection I shall present empirical evidence that supports the following generalization: null referential subjects are not allowed in Capeverdean matrix clauses.
In the next set of matrix clauses the subject content could be easily recovered from the context, given the deictic type of reference involved and/or other relevant information given in the sentence. Nevertheless, a null subject is ruled out.

Before looking at these relevant data, both in matrix and embedded clauses, let us recover the paradigm of the pronominal forms in the language, as in (183), here as (270).

(270) Capeverdean pronominal forms

<table>
<thead>
<tr>
<th></th>
<th>Emphatic forms</th>
<th>Free forms</th>
<th>Subject clitics</th>
<th>Object clitics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg.</td>
<td>ami</td>
<td>mi</td>
<td>N</td>
<td>-m</td>
</tr>
<tr>
<td>2sg (informal)</td>
<td>abo</td>
<td>bo</td>
<td>bu</td>
<td>-bu/-u</td>
</tr>
<tr>
<td>2sg (form, masc)</td>
<td>anho</td>
<td>nho</td>
<td>nho</td>
<td></td>
</tr>
<tr>
<td>2sg (form, fem.)</td>
<td>anha</td>
<td>nha</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3sg (fem., masc)</td>
<td>ael</td>
<td>el</td>
<td>e</td>
<td>-l</td>
</tr>
<tr>
<td>1pl</td>
<td>anos</td>
<td>nos</td>
<td>nu</td>
<td>-nu</td>
</tr>
<tr>
<td>2pl</td>
<td>anhos</td>
<td>nhos</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3pl</td>
<td>aes</td>
<td>es</td>
<td></td>
<td>-s</td>
</tr>
</tbody>
</table>

Some of the aforementioned crucial data follows:

(271) *Bu kre bai ku mi? *(N) ta bai mar.*

2SG want go with me 1SG TMA go sea.
‘Do you want to go with me? I’m going to the beach.’

(272) *(N) ta bai. Ti manha.*

1SG TMA go PREP tomorrow
‘I’m leaving. See you tomorrow.’

(273) *(N) txoma-s p’-es ben nkontra ku mi.*

1SG call-3PL TMA PREP-3PL come meet PREP 1SG
‘I called them so that they come join me.’
(274)  Manel, undi ki *(bu) sta?
    Manel where that 2SG be
    ‘Manel, where are you?’

(275)  Obi li, oxi me *(nhu) ta traze-m karneru!
    listen here ADV 2SG TMA bring-1SG sheep
    ‘Listen, you bring me the sheep today.’

(276)  Ka *(nhos) bai dja!
    NEG 2PL go already
    ‘Don’t you leave so soon!’

Baptista 2002’s argumentation in favor of the pro-drop status of the language is
grounded on the observation of the following data:

(277)  (El)   e nha pai.    (Baptista 2002:255)
    He COP my father
    ‘He is my father.’

(278)  a. Bu sta livri.
    you are free
b. sta livri.
    is/are free

From these sentences with an optional subject – which are not common at all, and
whose properties seem restricted to rather particular contexts –, the author concludes
that null referential subjects are possible with some types of predicates.

I have argued before (in Pratas 2002, 2004b) that cases of 3SG as in (277) can be
analyzed as phonological coalescence – clitic e + copula e. Otherwise it would be
difficult to explain why (279b) is not possible, for instance, with the 1SG reading:

(279)  a. E nha pai.
    3SG+be my father.
    ‘He is my father.’
b. *E spertu.

BE clever

‘He/she is clever.’ / * ‘I am clever.’

However, as said in Chapter Three, this coalescence hypothesis does not seem a right proposal either, since the copula prevents the clitic form from occurring in the subject position, requiring the free form to apply:

(280)  

a. * Ne spertu.

b. Mi e spertu.

In other words, the examples in (278) could not be derived from something like we have in (281), since these sentences are ungrammatical. The well-formed original sequence available would be (281c):

(281)  

a. * E e nha pai.

b. *E e spertu.

c. El e spertu.

The Present copula e is the only Capeverdean verb form with this effect on the subject pronominal selection, and the reasons for this are not quite clear. The fact that only this form – and not the Past form era – appears in a pre-negation position, as we have seen earlier, might have some relation to this restriction, but this must be subject to further analysis.

In any case, what is under discussion here is that when a first person singular is desired the sentence in (280b) is in fact the only well-formed counterpart for (279b) or (281c).

The aforementioned impossible 1SG reading for (279b) could be due to the fact that, while “null subjects of individual-level predicates are recoverable as 3rd person singular argumental pronouns el or e, null subjects of stage-level predicates may be interpreted as 1st, 2nd or 3rd person.” (Baptista 2002:256) This would be what we had in (278b).

If this is so, however, the impossibility of null subjects with a stage-level predicate in simple declarative contexts is rather unexpected:
   ‘Am sick.’
   b. N sta duenti.

Given this contrast, perhaps the exceptional case in (278b) is that the null subject is licensed by a special idiosyncrasy of the exclamation context, which in turn is dependent on the semantics of the predicative construction in question: the announcement that someone is (finally) free.

As for the possibly absent subject in (277), given these arguments there is a much more coherent possibility. Given the impossibility of null referential subjects except for these very specific contexts (3SG + present tense copula), I have argued earlier in favor of some kind of haplology: the 3SG clitic and the copula, both e, coalesce phonologically as a single e.

Still a different possibility seems available. Consider a discursive environment with a man entering the room; someone says E nha pai ‘Is my father’. This null element may be a variable, bound by a given context information. Two observations seem to favor this proposal:

i) as expected from what happens with some other contexts where a bound variable is possible in Capeverdean, this is not possible in English;
ii) as it happens with the contexts I will analyze in the next subsection, this bound variable is only possible with a 3SG reading.

This hypothesis has, however, some flaws; in particular why this discourse-bound variable is not available with other verbs.

Independent evidence has led me to consider a different possibility: the pro here is not referential, but expletive. If a sentence like ‘It’s my father’ is possible in English with an overt expletive subject, it might be possible in Capeverdean with an expletive pro. My present approach on expletive environments, however, could motivate still a simpler view on these contexts: this is in fact an expletive construction; hence nothing is in subject position.
Either way, the crucial statement at stake in this section is that I have not found empirical evidence for the claim that Capeverdean is a pro-drop language. The dependency contexts that will be presented in the next subsection do not challenge this statement. The reasons for this will be completely understood in section 5.2. For the moment, let me describe these particular embedded null subjects.

### 5.1.2 Some dependency contexts

In most embedded full clauses – I am referring to the whole set of the clearly finite clauses which are introduced by the complementizer *ma* (thus, these are embedded CPs) – null subjects are ruled out as well. This prohibition occurs, again, even when the subject content is easily recovered from the context:

(283) a. Madalena *fla-m m’-e ka konxe kel livru nobu di Germano Almeida.*

   Madalena say-1SG COMP-3SG NEG know the new book by Germano Almeida
   ‘Madalena told me she does not know the new book by Germano Almeida.’

b. * Madalena *fla ma ka konxe.*

However, and as discussed in Costa & Pratas 2006, when some embedded contexts cross with different types of matrix subjects, null referential subjects emerge.

This null subject may be in an embedded CP:

(284) a. Kenha *ki ta atxa m-e spertu labanta mo.* (who=embedded subject)

   Who REL TMA find COMP-BE clever raise hand
   ‘Who finds/considers himself clever, please raise your hand.’

b. Kenha *ki ta atxa ma el e spertu labanta mo.* (who≠embedded subject)

   Who REL TMA find COMP 3SG is clever raise hand
   ‘Who finds that he/she is clever raise your hand.’

Notice that any 3sg clitic could not have been subject to any phonological operation because of the Present copula since, as shown earlier, with this the obligatory pronominal form is what I have called the free form – which here behaves as an X°.
In order to discard this hypothesis we may try this with a different verb:

\[(285)\]

a. *Kenha ki ta atxa ma sabe gia kel karu-li?* (who=embedded subject)

Who REL TMA find COMP knows drive REL car-LOC

‘Who finds/considers himself able to drive this car?’

b. *Kenha ki ta atxa ma e sabe gia kel karu-li?* (who≠embedd. subject)

Who REL TMA find COMP-3SG knows drive REL car-LOC

‘Who finds/considers that he/she is able to drive this car?’

And one more:

\[(286)\]

a. *Kenha ki fla ma gosta di papaia?* (who=embedded subject)

Who REL say COMP like papaya

‘Who said to like papaya?’

b. *Kenha ki fla ma e gosta di papaia?* (who≠embedd. subject)

Who REL say COMP-3SG like papaya

‘Who said that he/she likes papaya?’

But this null subject may likewise be within a relative clause embedded in a DP:

\[(287)\]

a. *Ningen ka ta atxa livru ki perdeba.* (nobody=emb. sub.)

Nobody NEG TMA find book REL loose.TMA

‘Nobody finds the book that they (each of them) had lost.’

b. *Ningen ka ta atxa livru ki-e perdeba.* (nobody≠emb. sub.)

Nobody NEG TMA find book REL-3SG loose.TMA

‘Nobody finds the book that he/she had lost.’

It has been noted in Rizzi 1986, Sigurðsson 1993, Coelho et al. 2000, Kato 2002, among others, that the null subject parameter is not a uniform phenomenon; it depends on:

i) the availability of null forms and

ii) the licensing conditions for null forms.
Following this idea, in Costa & Pratas 2006 the analysis for Capeverdean goes as follows:

i) \textit{pro} is available;

ii) \textit{pro} is a bound-variable.

Hence, with respect to the licensing condition in ii), Capeverdean differs from regular \textit{pro}-drop languages, in which \textit{pro} is licensed by Agr.

This proposal is supported by a comparative approach to Brazilian Portuguese, which some authors argue is undergoing a change regarding the status of the null subject parameter (Duarte 1995, a.o.). According to Duarte (1995), we can observe a massive decrease of null subjects in a diachronic study on popular plays written in the 19\textsuperscript{th} and 20\textsuperscript{th} century: 1845 – 20\%; 1882 – 23\%; 1937 – 46\%; 1975- 67\%; 1992- 74\%.

Nevertheless, as noted in Figueiredo Silva 1996, among others, it cannot be said that Brazilian Portuguese is no longer a null subject language. Null subjects are found in several instances:

As examples of null subjects in Brazilian Portuguese we have question-answer pairs:

(288) Q: \textit{O que você comprou?} (what you bought)

‘What did you buy?’

A: \_ comprei um livro.

‘I bought a book.’

But we have null subjects in embedded contexts as well (despite the loss of the Avoid Pronoun Principle (Chomsky 1981)):

(289) \textit{O Pedro disse que (ele) ganhou na loto.} (from Modesto 2000)

‘Pedro said that he won the lottery.’
Modesto 2000 shows that the null interpretation of the null subject in (289) is limited. It is also different from what happens in a typical pro-drop language like European Portuguese:

(290) a. Brazilian Portuguese:

\[ O \ Pedro, \ disse \ que \ ec_{ij} \ ganhou \ na \ loto. \]  
‘Pedro said that he won the lottery.’

b. European Portuguese:

\[ O \ Pedro, \ disse \ que \ ec_{ij} \ ganhou \ na \ lotaria. \]

Crucially, in Brazilian Portuguese pro is necessarily bound by the c-commanding antecedent. This is similar to the contexts in which null subjects are available in Capeverdean. This has inspired the proposal in Costa & Pratas 2006, according to which in Capeverdean the null subject is bound by a c-commanding antecedent as well.

The difference between the two languages is the following:

i) in Brazilian Portuguese the antecedent need not be quantificational – it may be a DP such as \( O \ Pedro; \)

ii) in Capeverdean pro is only licensed if the antecedent is a specific kind of operator, a quantifier or a \( wh- \) expression.

As for the embedded pronouns, the two languages also behave differently:

i) in Brazilian Portuguese the embedded overt pronoun may be co-referential with the matrix subject;

ii) in Capeverdean this is not possible: with certain operators in the matrix subject position, thus c-commanding the pronominal embedded subject, co-reference between both subjects is excluded (as we have seen in the b. versions of examples above – (284)-287)).

In Modesto 2000, it is proposed that in Brazilian Portuguese all subjects may occupy an A-bar position, from which they can bind a variable. As shown by Modesto 2000, Costa & Galves 2002, and Costa, Duarte & Silva (in press), this does not mean that the subject occupies a peripheral position – the A-bar status may derive from specific properties of the Brazilian Portuguese extended IP domain.
In Capeverdean, as shown in Chapter Two, DP subjects may occupy an unsplit Spec,TP, an A-position, but also a peripheral position in clitic doubling contexts (recall that, although this left-dislocation is not so obvious in sentences like Ami-N ta bai sinema ‘I’m going to the movies’, we have more clear topicalization contexts like Djon, el gosta di filmi ‘Djon, he liked the movie’).

The position of regular DP subjects, however, is not relevant for these Capeverdean operator-variable contexts, since, as we have just seen, the binding operators can only be either a wh- or a quantified expression (which may be argued to have an inherent operator status, no matter what their position is). Either way, for the wh-operator, these predictions seem correct.

But what can we know about the quantified operators? In the examples above, we have a null embedded subject when the matrix subject, the operator, is the negative ningen ‘nobody’. But what if we have the universal quantifier tudu algen ‘everybody’? The pattern is maintained:

(291) a. Tudu algen atxa livru ki perdeba. (everybody=embedded subject)
   Everybody find book REL loose.TMA
   b. Tudu algen atxa livru ki-e / es perdeba. (everybody≠embedded subject)
   Everybody find book REL-3SG / 3PL loose.TMA
   ‘Everybody found the book that he/she/they had lost.’

(292) a. Tudu algen ta fla kel ki kre. (everybody=embedded subject)
   Everybody say DEM REL want
   b. Tudu algen ta fla kel ki-e / es kre. (everybody≠embedded subject)
   Everybody say DEM REL-3SG / 3PL want
   ‘Everybody says what he/she/they want.’

This operator-variable relation is also available when the DP dominating the relative is embedded in an adjunct PP (since this is a lower adjunct, c-command holds between the two positions).

(293) a. Ningen ka sta kuntenti ku kel ki ten. (nobody= subject of relative)
   Everybody NEG BE happy with DEM REL have
   ‘No one is happy with what one has.’ / * ‘Nobody is happy with what has.’
Notwithstanding that, the prohibition against the co-reference whenever there is an embedded pronoun seems not to hold.

In this section I have shown crucial empirical support for the view that Capeverdean is not a pro-drop language: it allows no referential null subjects in matrix clauses and, in the rare embedded context where null subjects are permitted, these may be analyzed as a bound variable (Costa & Pratas 2006).

Relevant data both in matrix clauses (5.1.1) and embedded clauses (5.1.2) have been presented, along with the lines of reasoning that lead to the observation that these do not support the pro-drop status of the language, with the respective descriptions and proposed analyses.

As for the lack of overt expletives in Capeverdean, I now propose a different approach: there are no null expletives either. I will approach it in section 5.3. This proposal is accompanied by the intuition that these constructions should not be taken as some sort of property involved in the pro-drop parameter.

The relevant properties for considering a language as pro-drop will be discussed in the next section. Following Rizzi 1986:

i) overt agreement morphology on the verb has been considered necessary to provide the content of the silent category pro, both in referential and expletive constructions;

ii) as for the formal licensing of the null subject, it is required that it is in a specifier-head relation with the verb.
Besides the broad debate about what might be considered as rich morphology, we shall see that there is more to the null subject parameter than these two conditions.

5.2 The null subject parameter

As said earlier, some languages allow for the subject not to be pronounced, once it is in the specifier position of the relevant functional projection, say in Spec,IP. This position is then filled with a silent category, pro. At least for Romance languages like Italian and European Portuguese, this must happen after the subject moves to this position, where it is argued the DP checks its Case features.

On the other hand, as the verb moves to I, the agreement morphology that is affixed onto the verb is also (besides being a motivation for V-to-I, as discussed earlier) one of the properties responsible for the licensing of pro, for information on person and number provide the content of the null subject (according to the hypothesis in Rizzi 1986, the relevant agreement morphology for this purpose is the one for person, not for number). This syntactic configuration, the subject on Spec,IP and the verb on I, guarantees the specifier-head relation that constitutes the other requirement, strictly structural, for a language to allow pro.

Following this view, the crucial point for licensing null referential subjects in many languages seems to be agreement morphology, which must be sufficiently “rich”, given that the spec-head relation depends on this as well. This is what is lacking in English (only the third person singular morpheme in the present tense, no person or number morphemes in other tenses) and also in French (despite the apparent person and number distinctions in the written form).

At this point, a question emerges: what might be considered “rich morphology”, or, in the words in Roberts 1999, a “relevant morphology”? Capeverdean is totally out of this debate176, since it has no overt agreement morphology at all. This is visible in the verbal paradigm in (295), which applies to all classes of predicates: unergative verbs,

176 Out of the debate on what the relevant morphology in the licensing of null subjects might be, not the debate on the null subjects itself – the fact that Capeverdean has no overt agreement morphology, per se, does not preclude it from having null referential subjects, since, as noted in Huang 1984, Chinese does have null subjects and also no overt agreement morphology (in this case, it has been argued that null subjects are licensed by some discourse related mechanism).
like *tilifona ‘call’, transitive verbs like *odja ‘see’ and unaccusative verbs, like *txiga ‘arrive’, in all tenses (although only three are illustrated, with *kanta ‘sing’).

(295) Lack of overt verbal agreement morphology in different tenses: *kanta ‘sing’

<table>
<thead>
<tr>
<th></th>
<th>PastPer</th>
<th>Pres</th>
<th>PastHab</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>*N kanta</td>
<td>*N ta kanta</td>
<td>*N ta kantaba</td>
</tr>
<tr>
<td>2SG</td>
<td>*Bu kanta</td>
<td>*Bu ta kanta</td>
<td>*Bu ta kantaba</td>
</tr>
<tr>
<td>3SG</td>
<td>*E kanta</td>
<td>*E kanta</td>
<td>*E ta kantaba</td>
</tr>
<tr>
<td>1 PL</td>
<td>*Nu kanta</td>
<td>*Nu kanta</td>
<td>*Nu ta kantaba</td>
</tr>
<tr>
<td>2 PL</td>
<td>*Nhos kanta</td>
<td>*Nhos kanta</td>
<td>*Nhos ta kantaba</td>
</tr>
<tr>
<td>3 PL</td>
<td>*Es kanta</td>
<td>*Es kanta</td>
<td>*Es ta kantaba</td>
</tr>
</tbody>
</table>

In 3.2.1, when the verb movement was at issue, some questions around the role of overt verbal agreement morphology have been called into discussion. In fact, the markings for person, number and gender have been frequently involved in the debate about some other parametric values – besides verb movement and the licensing of *pro.

In Castro & Pratas 2006, a possible relation between some different parameters has been addressed – not conclusively, though. The main topic of that paper has been the reduced agreement morphology within the Capeverdean DP, as illustrated:

(296)  a. *livrus bunitu
   book-PL beautiful

   b. *purmerus livru
   first-PL book

(297)  a. *uns livru bunitu
   a-PL book beautiful

   b. *alguns livru bunitu
   some-PL book beautiful
c. kes livru bunitu
these-PL book beautiful

d. nhas livru bunitu
my-PL book beautiful

Following the line of reasoning in Enç 1991, Costa & Figueiredo Silva 2003 argue that in Brazilian Portuguese the plural morpheme is realized as a singleton on D, the head that links the DP to its LF interpretation.

Taking this as a departure, the proposal in Castro & Pratas 2006 goes as follows: in the Capeverdean DP, the plural marker is a singleton and it surfaces on the D head (as a suffix): if there is an overt element on D, it attaches to that element; if there is no overt element on D, it “lowers” to the head of the complement - as in (296), with plural surfacing on the noun and on the adjective.

The relevance of this point to what is at issue here is that in Castro & Pratas 2006 it has also been proposed there could be some relation in Capeverdean between the reduced agreement morphology within the DP and the lack of overt verbal agreement.

As said earlier (3.2.1), Roberts 1999 departs from Chomsky 1993’s proposal according to which the relevant linguistic parameter concerns the value of an abstract feature that licenses verbs and is associated with I. This feature is called I’s V-feature and it varies parametrically as either strong or weak. Roberts 1999 relates the value of the parameter with the richness of agreement morphology:

- If there is verbal agreement marking of the relevant type, then I has a strong V-feature.

At this point, we are again dealing with the crucial question: what is the “relevant type” of morphology? Roberts 1999 analyzes the relation between the presence of agreement morphology, particularly plural endings, and the triggering of the strong V-feature. In a previous analysis, in Roberts 1992, the author had suggested that overt distinct marking for number is the relevant kind of “richness”. For Rohrbacher 1994, though, Roberts makes the wrong predictions for Faroese, a language that lacks verb movement but has overt, distinct plural marking. Instead, Rohrbacher 1994 proposes that:
Distinct first- and second-person markings in at least one number of one tense of the regular verbs are required.\textsuperscript{177}

On the other hand, Vikner 1995’s generalization predicts that:

- A language has V-to-I movement if there is Inflection for Person in all tenses.

A brief overview on agreement paradigms in some languages is presented in (298), extending the data in Costa & Figueiredo Silva 2003. To the languages under analysis by these authors (European Portuguese and two varieties of Brazilian Portuguese, BP1 and BP2), Capeverdean and English have been added.

(298) Verbal agreement paradigms

<table>
<thead>
<tr>
<th></th>
<th>EP</th>
<th>BP1</th>
<th>BP2</th>
<th>English</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Singular</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>como</td>
<td>como</td>
<td>como</td>
<td>eat</td>
<td>kume</td>
</tr>
<tr>
<td>2</td>
<td>comes</td>
<td>come</td>
<td>come</td>
<td>eat</td>
<td>kume</td>
</tr>
<tr>
<td>3</td>
<td>come</td>
<td>come</td>
<td>come</td>
<td>eat</td>
<td>kume</td>
</tr>
<tr>
<td><strong>Plural</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>comemos</td>
<td>come/comemos</td>
<td>come</td>
<td>eat</td>
<td>kume</td>
</tr>
<tr>
<td>2</td>
<td>comem</td>
<td>comem</td>
<td>comem</td>
<td>eat</td>
<td>kume</td>
</tr>
<tr>
<td>3</td>
<td>comem</td>
<td>comem</td>
<td>comem</td>
<td>eat</td>
<td>kume</td>
</tr>
</tbody>
</table>

In (298) we observe that, in colloquial EP, there is a complex paradigm, with only two overlapping forms, but clearly an “overt, distinct, plural marking”.

On the other hand, in BP1, despite a more simplified paradigm, number is a distinctive feature. Still, it is crucial to observe that there are no distinct first- or second-person markings – and we assumed then that BP1 lacks them in all tenses of regular verbs: first person has a distinct marking but second person is not distinct from third person.

The same applies to BP2, English and Capeverdean. In BP2 and in English, there is a much more simplified paradigm, with only one distinct form (first person singular in BP2 and third person singular in English) and no plural distinction. In

\textsuperscript{177} This criterion covers Faroese, which, unlike Icelandic, lacks such marking – it is, thus, expected that it also lacks verb movement.
Capeverdean, there is no distinction at all. Therefore, these data are sufficient to
determine the value of the proposed I’s V-feature in these languages, as we can see in
(299). Notice that what is called I is the relevant functional head in each language. It can
be AgrS, such as in EP, or T, such as in Capeverdean.

(299) Some parametric values + DP-internal number agreement

<table>
<thead>
<tr>
<th></th>
<th>EP</th>
<th>BP1</th>
<th>BP2</th>
<th>English</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevant verbal morphology (as in Rohrbacher 1994)</td>
<td>+</td>
<td>-178</td>
<td>-</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>V-features of I (AgrS or T)</td>
<td>strong</td>
<td>weak</td>
<td>weak</td>
<td>weak</td>
<td>weak</td>
</tr>
<tr>
<td>V-to-I movement</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Referential pro</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DP-internal number agreement</td>
<td>+</td>
<td>(300)</td>
<td>(300)</td>
<td>(300)</td>
<td>-</td>
</tr>
</tbody>
</table>

Notice that the lack of referential pro in both varieties of Brazilian Portuguese refers to
matrix clauses, in no context of question/answer pair. As it has been illustrated in 5.1,
following Modesto 2000, there are specific embedded contexts where a null referential
subject is allowed. This has been analyzed as a bound variable, hence it is licensed in a
different way – in other words, it is distinct from the null subjects dependent on
agreement morphology and possibly related to other parametric values.

In (299) the values for the relevant verbal morphology, as proposed by Rohrbacher
1994, are included, and likewise the V-features of I, V-to-I movement, referential pro
and DP-internal number agreement. The last line of (299) means: while the DP-internal
number agreement is strong (+) in European Portuguese, with the relevant morpheme
being affixed to all words, and while in Capeverdean this agreement is weak (-), with
the relevant morpheme occurring only once, the two varieties of BP considered in Costa
& Figueiredo Silva 2003 and also English represent some different forms of a
weakened/impoverished paradigm.

This contrast can be observed in (300).

178 In spite of BP1 having number distinctions (in what it differs from BP2), the relevant property is
having distinct first and second person markings, and for these ones the value is minus.
Expletive *pro* is not included in (299) and this omission has different possible explanations. One of them, sufficient for the moment, follows from the claim that “expletive null subjects are a lexical, rather than a parametric, option [...] its distribution varies depending on the other expletive pronouns available in the language.” Roberts (1999:313). Accordingly, expletive *pro* is not related to the V-features of I. As we shall see in the next section, a better reason for expletive *pro* not to be included here is that in Capeverdean there is, simply, no expletive *pro*.

We can observe, from (299), that BP1 and mainly BP2 show that the relation between verbal morphology, strong V-features of I and then verb movement is a one-way implication. In spite of having no relevant verbal morphology, BP1 and BP2 still have verb movement, as proposed in Costa & Galves 2002. These authors claim that there is no evidence to posit different landing sites for the verb in EP and BP (both in BP1 and BP2), and, with Costa 1996, that there is verb movement from V to T, without reaching the topmost functional head. Therefore, BP1 and BP2 seem to illustrate the claim that “the loss of verbal agreement marking on its own is not sufficient for the change in the value of I’s V-feature, and hence, such a change on its own has no effect on V-to-I movement.” (Roberts 1999:292)\(^{179}\)

Now consider for Capeverdean the relation between the V-features of T and the conditions for licensing of referential *pro*. The raising of subjects from Spec,VP to Spec,TP is due, still according to Roberts 1999, to a property of UG (not a parametrically variant property): the strong N-features of the respective head, AgrS.

Suppose we assume, still according to Roberts (1999:309), that the conditions for referential *pro* to be licensed are:

\(^{179}\) As mentioned in 3.2.1, it is attested in the literature and in diachronic analyses (for English, Swedish and Danish, among other languages) that the loss of a rich verbal morphology preceded, sometimes in one hundred years, the loss of verb-movement – Kroch 1989, Lightfoot 1991, Platzack 1987. Consequently, there may be languages that have verb movement but lack the relevant type of inflectional morphology. This seems to be the case of BP1 and BP2 – the properties of these two languages can be analyzed in terms of language change.
i) *pro* must be formally licensed by a strong N-feature and

ii) *pro’s* content must be identified by Spec-head agreement with the relevant inflection.

Considering that Capeverdean satisfies i) but not ii), one must conclude that *pro* is ruled out.

Let me explain the steps involved in this. As demonstrated in Chapter Two, there is no empirically grounded reason to posit a Split TP for Capeverdean (hence, there is no AgrSP node), but consider that these strong N-features, as a universal, are somehow present in the TP functional node. Referential null-subject languages where *pro’s* content is recovered morphologically (not, again, the case of Chinese, for instance – see Huang 1984) always have V-movement to the functional head.

As we have just seen, the converse does not necessarily hold: languages with V-movement do not necessarily allow null subjects, as Middle English, French or BP. In other words, referential null subjects of this type, whose content is morphologically recovered, are a marked property contingent to V-movement. In Capeverdean, V does not move to T; therefore, any potential *pro* would never be in Spec-head relation with the relevant inflection. Moreover, there is no such thing as “relevant inflection” in the language. We must conclude, thus, that it is not at all unexpected that matrix *pro*, as licensed by any type of overt agreement, is ruled out. As illustrated in 5.1, the only null referential subject allowed is some type of embedded *pro* that is licensed by a different mechanism – it is a variable bound by an operator.

All these conditions and assumptions seem to guide us to a relation, in Capeverdean, between verbal agreement, which is the independent parametric value, and what happens within the DP (this relation has been proposed in Castro & Pratas 2006): the plural surfaces post-syntactically, in the morphological component, as a singleton morpheme.

The V-features of T are weak, within the DP-subject surfacing on Spec,TP some minimal plural marker is required for the correct interpretation at LF. Assuming that D is the locus of plural marking, linking the DP to its LF-interpretation, thus the minimal marker surfaces on it.
Therefore, there seems to be a one way relation between weak DP-internal agreement and weak verbal morphology, since none of the languages that shows weak I’s V-features (BP1, BP2, English, CV) shows full DP-internal agreement. Exactly with which property it is related is still an open question.

At this point, we are led to a conclusion following the claim in Sigurðsson 1993: null subjects can be of different types. The key factor for identifying a pro-drop language is, then, not the availability of pro, but the way it is licensed. Consider the list in (301):

(301) Types of null subjects:
   a) Expletive (e.g. German)
   b) Topic-drop (e.g. diary English)
   c) Controlled by a c-commanding antecedent (e.g. Hebrew)
   d) Agr-related null subjects (pro-drop)(e.g. Italian, European Portuguese)

Hence, the only pro-drop languages are the ones in which null referential subjects in matrix clauses depend on licensing mechanisms connected to agreement morphology (type (301d). Capeverdean rare null subjects are akin to the embedded variables in Brazilian Portuguese referred to in 5.1, as contended in Modesto 2000. And even with respect to these Brazilian Portuguese bound variables, there is an important difference (whose consequences are left for further research): whereas these null embedded subject in Brazilian Portuguese can be bound by a matrix DP, Capeverdean variable can only be bound by wh- or a quantifier expression.

In this section I have presented some arguments available in the literature on the relevance of verbal agreement morphology both for the trigger of V-to-I and for the licensing of null subjects. Since these parameters are related to each other under these views, in Castro & Pratas 2006 a connection between these and a different parametric value – the person and number agreement within the DP – has been discussed.

Given that Capeverdean has neither overt verbal agreement morphology nor V-to-T, the type of null subject, pro, that is licensed under agreement and also a Spec-head relation with the verb is ruled out. The conclusion is: as shown in 5.1, the available null referential subjects in the language (embedded variables bound by an operator, akin to
some cases in Brazilian Portuguese), are of a different nature. Hence, Capeverdean is certainly not a pro-drop language.\footnote{I hope this discussion has also some effect on the future research on the null subject phenomenon, namely on the increasing number of acute questions concerning this label. As Maria Lobo points out, the interesting line of inquiry is more focused on the extensive description of the contexts and conditions in which different languages allow null subjects, and less on the label pro-drop.}

In the next section the so-called null expletives will be finally at issue. Capeverdean expletive constructions, where no overt subject is realized, may provide evidence in favor of a parametric variation of the EPP – with a negative value in Capeverdean. In other words, since there is no semantic subject in these sentences, and also no expression empirically attested as obligatory in subject position, why should we not consider that Spec,TP is not projected? I hope that in the end of this chapter an answer to this question has become clear.

### 5.3 The special case of null expletives

In this last section of the present chapter expletive constructions are analyzed. The main claim will be that no null subject – no expletive pro – occupies the subject position in these contexts. In other words, Spec,TP, which is not needed, is not projected.

The data in subsection 5.3.1 will cover both the cases where there is no subject at all (weather verbs, existential contexts, extraposition), and also some cases where the subject is allowed not to be in the subject position (unaccusative constructions with indefinite subjects and every other case where the subject is only a clitic – hence, it surfaces on T, despite its argumental status, and nothing is in Spec,TP).

In subsection 5.3.2 the EPP, the only theoretical requirement that might oppose to this empirical evidence, will be discussed and argued to be subject to cross-linguistic variation. In Capeverdean, it does not hold. Or, in other terms, the value of the EPP parameter is negative.

#### 5.3.1 When no subject is there

In Pratas 2002 I have assumed that, despite having no null referential subjects, Capeverdean allowed for all types of null expletive subjects. In subsequent works I have
argued that expletive null subjects do not correlate with verbal agreement (contrarily to what had been proposed in Rizzi 1986, which predicted that, also related to verbal morphology, the null expletive could always receive the 3sg interpretation). In most recent works, a more accurate knowledge about the language grammar has led me to the conclusion that Capeverdean has no null expletives at all.

Consider the following examples.

Existential constructions\(^{181}\)

With *ten* ‘have’, in all TMA contexts:

(302)  
\[
\text{TMA have one match of soccer important day Sunday REL TMA come.}
\]

‘There is going to be an important soccer match next Sunday.’

(303)  
\[
\text{have-TMA here one man REL TMA play-TMA drum}
\]

‘There used to be a man here who played the drum.’

(304)  
\[
\text{have coffin REL BE more cheap}
\]

‘There are coffins that are cheaper.’

With verbs like *fika* ‘stay’ or *resta* ‘remain’:

(305)  
\[
\text{If.3PL buy DEM more cheap TMA remain money \ PREP family REL TMA stay}
\]

‘If they buy the cheaper one there will be money left to the surviving family.’

(306)  
\[
\text{Patricia, TMA stay only one piece}
\]

‘Patricia, there is only one piece left!’

\(^{181}\) In English, the overt expletives obligatory in existential constructions are of the *there*-type, as opposed to the *it*-type. The *it*-type expletives are considered quasi-argumental and the *there*-type are the so-called ‘dummy subjects’. In Capeverdean, though, these labels are not relevant.
(307)  *Lisa, spera, ta falta so un grodi.*  
Lisa, wait, TMA miss only one braid  
‘Lisa, wait, there’s only one braid left!’

**Weather verbs**

(308)  *Pedru fla-m ma sata txobe na Lisboa.*  
Pedru say-1SG COMP TMA rain in Lisbon  
‘Pedru said it is raining in Lisbon.’

**Parse ‘seem’ constructions (raising)**

(309)  *Parse-m ma sata txobi na Lisboa.*  
seem-1SG COMP TMA rain in Lisbon  
‘It seems that it is raining in Lisbon.’

(310)  *Ta parseba ma ningen ka ta staba na kasa.*  
TMA seem-TMA COMP nobody NEG TMA COP-TMA in house  
‘It seemed that nobody was in the house.’

**Extraposition / adjectival predicates**

(311)  *E kapas ki txuba ben oxi.*  
BE able COMP rain come today  
‘It is likely that it will rain today.’

---

182 I cannot refer to these as ‘raising’, since in Capeverdean the subject does not actually raise. In other words, these are always constructions with no matrix subject. Observe the following contrast:

(i)  a. *Ta parse-m (gosi) ma Djon kre Maria txeu.*  
TMA seem-1SG NOW COMP Djon kre Maria a lot  
‘It seems to me (now) that Djon loves Maria.’

b. *Djon ta parse-m ma kre Maria txeu.*

(ii)  *Djon, ta parse-m m’-e kre Maria txeu.*  
Djon TMA seem-1SG COMP-3SG like Maria a lot  
‘Djon, it seems to me that he loves Maria.’

The sentence in (ii) is possible when certain discourse contexts demand topicalization. For instance, someone is talking about Peter, and who he likes, and then someone introduces the subject Djon, like saying “as for Djon…”.
(312)  *E faxi pa bu prendi badja tornu.*  
    BE easy COMP 2SG learn dance tornu  
    ‘It is easy for you to learn how to dance tornu.’

Clefts

(313)  *E di dinheru ki N sata mesti.*  
    BE of money COMP 1SG TMA need  
    ‘It is money that I need.’

As noted earlier, English always need an overt subject in sentences like in (314).

(314)  a. *(It) is raining.
    b. *(It) seems that John is sick.
    c. *(It) is likely that John comes to your party.

We know that there is no x such that x rains (weather predicate), there is no x such that x seems that John is sick (raising construction), and there is no x such that x is likely that John comes to your party (extraposition). Why languages like English behave like this is one of the core questions raised by Generative Grammar.

On these pronouns, the it-type of the so-called expletive subjects, we can say that:

i)  they act as mere slot-filters;

ii) they present no semantic contribution to the sentence;

iii) they are elements in NP positions but no theta-role is assigned to them.

We do not know if they represent exactly the same category within raising constructions, with weather predicates or in extraposition contexts. If any Case feature is checked through them, it is not clear why and how.

There is a different type of expletive subjects in English, in existential sentences: the word *there*, considered the truly dummy subject – besides being φ-less (no features for person or number), it is certainly caseless. In contrast with this *there*-type of expletive subjects, the it-type is sometimes considered quasi-argumental. Unlike *it, there* is associated with NP-subjects which have been moved to the right in the sentence, and it
cannot be associated with clausal subjects in extraposition contexts, as we can see by
the ungrammaticality of the sentence in (315):

(315) * There is likely that John comes to your party.

As is easily observed, these expletive subjects, both of the *it* - and the *there*-type, are
obligatorily overt for some reason other than the internal (semantic) requirements of the
predicates in question, since in other languages the same verbs – which, we must recall,
do not have an external theta-role to assign – must occur without a pronounced word in
the subject position, as we observe in the Brazilian Portuguese sentences in (316),
corresponding to the English ones in (314).

(316)  a. (*Ele) Está chovendo.
       b. (*Ele) Parece que o João está doente.
       c. (*Ele) É provável que o João venha à tua festa.183

But let us meet some other Capeverdean contexts where there is no subject in its
canonical position, although some DP subject is present in the sentence. These are, as
we have seen earlier, unaccusative constructions with indefinite subjects, which allow
subject-verb inversion184.

(317)  a. *(dja) Txiga tres algen.
       (TMA) arrive three people
       ‘Three people have (just) arrived.’
       b. *(dja) Txiga Juau. / Juau *(dja) txiga.
       (TMA) arrive Juau / Juau (TMA) arrive
       ‘John just arrived.’

183 In some non-standard varieties of European Portuguese these sentences may occur with a pronounced
3sg expletive pronoun.

184 In fact, in Capeverdean there is actually no “inversion”, since both the verb and the DP stay in situ: the
subject is generated post-verbally. The only reason why the expression “inversion” makes sense is
because the language is said to be SV and here the order is VS.
(318) a. *Dja kai un abion.
TMA fall/crash IND plane
‘A plane just crashed.’

b. *Dja kai abion di Lisboa.
TMA crash plane from Lisbon

c. Abion di Lisboa dja kai.
plane from Lisbon TMA crash
‘The plane from Lisbon just crashed.’

The summary of different approaches on this follows.

As noted above, Brazilian Portuguese exhibits a pattern similar (although with some differences) to Capeverdean. These are the relevant facts.

(319) a. *Comeu a maçã.
ate the apple

b. Tá chovendo.
BE raining

c. Chegou um homem.
arrived a man

d. Parece que o João está doente.
seems COMP the João BE.PRES sick

e. Tem mais três pessoas vivendo na minha casa.
has more three people living in my house

Summing up, referential pro’s in finite clauses are ruled out in most contexts (the particular ones in which they are allowed are not exactly the same in both languages); no overt pronouns are allowed in expletive contexts (the same pattern as Capeverdean).

Figueiredo Silva 1996 has proposed that, in Brazilian Portuguese, null expletive subjects are licensed by number morphology, which is still systematically overt (although this language has lost some person morphology). Consider the paradigm presented for the variety previously called BP1.
In Capeverdean there is no number morphology either, hence – and once again – any relation of an expletive *pro* in the language with some type of morphological property is obscure.

Furthermore, this overt verbal agreement morphology argued to license expletive *pro* seems to be relevant to give it a third person singular interpretation. But why this identification is necessary, under a semantic point of view, also remains a mystery.

A different proposal has been presented in Roberts 1999, as exposed in the last subsection: expletive *pro* is not associated with referential *pro*; the latter is a marked parametric option, whereas the former is a matter of lexical option: when languages have overt expletive pronouns they must use them, and *pro* is ruled out. This does not seem to explain a lot, though.

In addition to this, how could we account for the case in Germanic languages, which exhibit overt expletives and also expletive *pro’s*? Their distribution depends on the structural position of the (supposed) subject. More specifically, here is the problematic fact (which Vikner 1995 assumes as an open question): when in Spec,IP expletives may be null in German, Yiddish and Icelandic, but, in the same position, they must be overt in Danish, Faroese and other Germanic languages.

Does this contrast constitute some extra empirical support favoring the fact that, in certain languages, Spec,IP is not projected if not needed?

Or could this be an important clue to be followed given that, in these Verb Second languages, C (and not T) is assumed as the position with Nominative Case properties? And how can this be related with, for instance, a Germanic Verb Second language like West Flemish, in which agreement on C licenses referential *pro*, but not expletive *pro* (from Cabredo Hofherr 1999).
K weten dat *(et) regent

I know that EXPL rains

'I know that it rains.'

The possible ways in which these phenomena relate to each other must be subject to further research and a careful analysis of different languages and expletive distribution patterns.

The line of reasoning that is determinant for now is that Capeverdean has no overt expletive of any type or in any type of environment. Given that it exhibits also no agreement morphology (even assuming that the so-called expletive pro is related to some marking for person, or number, or both), it is my proposal that it has no expletive pro either. In the cases where no lexical subject surfaces in the subject position, Spec,TP is not projected. In other words, this fact is not exclusive of expletive constructions, but occurs as well in environments whose subject never raises to Spec,TP: unaccusatives with indefinite subjects and clauses whose subject is a clitic, surfacing on T.

In the case that we consider abstract feature checking as a motivation for null entities, the subject position of the constructions mentioned seem eligible to apply the prediction in Chomsky 1995: feature checking should occur with no recourse to other functional projections that have no meaning and no phonological representation; that is, no interpretation at interface levels, LF and PF.

In this subsection I have addressed the disturbing question of which abstract rules of human language accounting for expletive subjects could account for the emptiness of these subject positions in some languages as well.

According to some previous analysis of different languages (Rizzi 1986, Figueiredo Silva 1986, Vikner 1995), the relevant properties are somehow related to agreement morphology and to some specific abstract properties of Spec,IP. In addition, Roberts 1999 contends that overt expletives and expletive pro are a matter of lexical variation: when languages have overt expletives, they use them, when they do not, they use the empty category. This does not account for Germanic languages which have both, with a distribution that depends on the structural position at stake.
Capeverdean presents a challenging association of facts:

i) no overt verbal agreement morphology in any tense;
ii) no overt expletives;
iii) no referential pro licensed by Agr.

This led me to the proposal that the way out of this puzzle must have something to do with the canonical subject position. In this language, whenever it is not needed in order to accommodate a lexical subject – some entity with both meaning and phonological representation (that is, interpreted at interface levels, LF and PF) –, this position is not projected. Recall that, under a minimalist perspective, it is movement that creates the specifier in the first place.

And here we are, back to the question above, now reformulated: which abstract rules accounting for obligatory expletive subjects in some languages could likewise account for the absence of the subject position (when not needed) in some others?

One possible answer is: the EPP. Notice that it can only hold as a solution to the puzzle as long as it is subject to parametric variation. This will be my final proposal in the next subsection, after reviewing some historical facts about this mysterious entity.

5.3.2 The EPP: a stone in the shoe

This subsection addresses one of the most enigmatic beings in the history of Generative Grammar: the Extended Projection Principle (EPP). It started as a generic and universal property about subjects and, after various adaptations, is now sometimes assumed as a property of features.

No matter what its precise nature is, some authors still surrender to its universal status, but others contend that it is not universal and others even propose that it does not exist. To follow this more radical view, though, it is necessary to propose an alternative approach to the facts in languages like English and French where, actually, the subject position must be filled with some phonological material. I will leave for another opportunity the analysis of the argumentation, for instance, in Epstein & Seely 2006. For the purposes of the present work, the crucial assumption will be that the EPP is, at least, not universal. More precisely, it is not active in Capeverdean.
The previous subsection provided empirical relevant facts on which this assumption is grounded. Now let us just recall some arguments around the EPP.

“The Extended Projection Principle (EPP) has been a pervasive mystery and a pervasive topic of research since it was first formulated by Chomsky 1981.” (Lasnik 2001:356)

The structural requirement that certain configurations have an obligatory subject applied at least to one class of languages, including English, even when the predicate had no θ-role to assign.

In Chomsky 1982 the modifier “extended” has been added to the original definition, after it has been admitted that the relevant requirement could not be inferred from the Projection Principle. The previous Projection Principle stated that “the θ-marking properties of each lexical item must be represented categorically at each syntactic level […]”. (Chomsky 1982:8) This implied that “subcategorized complements are obligatory for heads” but, as it can be seen from passives or nominals, “the θ-marked subject is not”. Thus, we cannot conclude that the fact that “clauses must have subjects” follows from the Projection Principle. It must follow from something else.

Here are the words: “[…] nonarguments can occupy the subject position, as in it is clear that S […] ; in fact, the subject position must be filled by a pleonastic element in structures lacking a θ-marked subject. It seems, then, that the requirement that a clause have a subject is independent of the Projection Principle […] I will refer to the Projection Principle along with the requirement that clauses have subjects as the Extended Projection Principle. (Chomsky 1982:9-10)

As Lasnik 2001 notes, the name EPP is assumed in the current terminology to refer only to the “extended” part of the Extended Projection Principle.

The EPP has also been part of the reformulations presented in the minimalist era, in particular in Chomsky 1995. Here it is viewed as a strong feature of a functional head high in the structure, causing the correspondent feature to raise.

This analysis of feature checking, combined with a PF-based generalized pied-piping requirement, gives way to the movement of the residual constituent, which, thus, also raises in order to check its feature.
In Chomsky 2000 there is a rejection of feature-based movement: what has been proposed as feature-checking via movement has now been replaced by a relation of long distance agreement, Agree. Hence, the EPP has no longer a feature-checking content, but, somehow returning to the previous view, “it is the requirement that certain functional heads must have a specifier.” (Lasnik 2001:357) In spite of considering the rejection of feature-based movement in favor of Agree as “too hasty”, Lasnik 2001 admits that, ironically, his line of reasoning leads to the conclusion that, just as Chomsky 2000 claims, “the EPP is not a matter of feature checking”.

As exposed in 3.2.1, after Chomsky 2001, and particularly in Pesetsky & Torrego 2001, 2005, the EPP is a sub-feature of a feature. Roughly, in their analysis of English T-to-C these authors establish a connection between this head-to-head movement and the presence of an unvalued uninterpretable T feature marked [+EPP] on C.

In sum, the EPP has been transformed from a requirement on clausal structure – all clauses have subjects – into a morpho-syntactic requirement of feature checking (there is a feature, an EPP feature, which attracts a DP to the specifier position of a functional head – similar to the general feature-checking mechanisms of Case and Agreement).

The EPP, in its different formulations, has been made responsible for some distinct phenomena. Sometimes these phenomena have been also accounted for by other explanations and principles – thus, the EPP has been accused of redundancy. Besides accounting for obligatory overt expletives in some languages, it also forced, for instance, successive cyclic movement (a category would have to stop in the intermediate Spec,TP on its way to the higher target position). To argue that this occurs because of the EPP is redundant with locality restrictions on movement.

---

185 Curiously, Noam Chomsky (p.c. 2004), says that the E, in EPP, should stand for exceptional.

186 But notice that, following Bošković 2002, expletives are always generated in their surface position – that is, they never move. Hence, it is not clear what their role is in this picture.
These accusations of redundancy have led Epstein & Seely 2006, among others\(^{187}\) (Martin 1999 had suggested it too), to the proposal that the EPP be eliminated; the so-called “EPP effects” would be, thus, deduced from independently motivated principles. Like the ones on Case/Agreement (with whose checking under spec-head the EPP overlaps), Theta theory, Locality conditions on Movement and the theory of null complementizers. Furthermore, the authors argue that “in the few cases where the EPP applies non-redundantly [on English], it is making the wrong predictions” (Epstein & Seely 2006:53)

One different proposal is the one in Landau 2005, defending a phonological approach to the EPP feature. The author follows Holmberg 2000 in calling it \([P]\), “a selectional feature that must be locally satisfied by some element with a phonological content.” In this view, null subjects can no longer satisfy this feature (Landau 2005 clarifies that selectional features are satisfied, uninterpretable agreement features are checked and deleted): this inability of satisfying this selectional feature holds for \textit{pro} (existent in Null Subject Languages, NSL), PRO or unpronounced movement copies (“traces”).

Let us review the author’s arguments regarding each of these null subjects.

Regarding \textit{pro}, Landau 2005 follows Alexiadou & Anagnostopoulou 1998: the function served by DPs and overt pronouns in non-NSLs is taken over by the inflectional morphology on the verb in NSLs – here, the finite verb overtly raises to T, checking all the features that an overt subject would check in non-NSLs. These features are D and Agr. The choice of an \(X^o\) (V-to-T) or an XP (DP-to-Spec, T) to satisfy \([P]\) on T would be parametric. Under these assumptions, “it is not even clear that \textit{pro} exists. If the inflectional morphology on the finite verb is interpretable, just like a pronoun, then \textit{pro} becomes redundant.” (Landau 2005:36)

In control contexts – and given that Landau assumes that PRO exists as a distinctive formative of the grammar\(^{188}\) – PRO (being a null category) cannot satisfy the EPP feature. Hence, either control clauses lack the EPP property or they have it but it is satisfied by something other than PRO, like inflection does in NSLs. For independent

---

\(^{187}\) Some authors have proposed explicitly the rejection of the EPP (Castillo, Drury & Grohmann 1999, Bošković 2002), others have presented analyses that are inconsistent with it (like Manzini & Roussou 2000). McCloskey 1996, on some Irish word order puzzles, casts some doubt on the EPP; namely through the formulation of the open question on which inflectional projection between C and V is privileged by it.

\(^{188}\) Cf. the discussion on this topic in 6.3.1.
reasons, Landau 2005 makes the option of considering that these constructions lack the EPP property. But, then, what makes PRO raise to Spec,TP if it does not satisfy the EPP? And what if PRO does not exist?

As for the unpronounced copies (traces) created by A-movement or A’-movement, the analysis in Landau 2005 regards selection in general: selectional requirements are satisfied upon merge in the selected position, before it is evacuated by movement. As presented in Landau 2005, just like in (322a) the moved constituent, \textit{which girl}, satisfies at LF the s-selectional requirement of V ([\textit{+animate}]), \textit{which noise}, in (322b) satisfies at PF the p-selectional requirement of T ([\textit{+P}]).\footnote{According to Landau 2005, s-selection governs LF configurations; p-selection is an interface condition which governs PF configurations; c-selection is the only type of selection that is part of narrow syntax.}

\begin{enumerate}
  \item a. Which girl, did that noise scare \textit{which girl}? \\
  \item b. Which noise, do you think \textit{which noise}, scared that girl?
\end{enumerate}

In this subsection I have reviewed the history of the Extended Projection Principle (EPP), which, with this “extended” definition in Chomsky 1982, assumed that “clauses must have subjects”. The definition has suffered adjustments and modifications, until the most recent minimalist view that the EPP is a property of a feature.

As applied to the TP node, it still means that Spec,TP is occupied; in other words – and in some cases – this is equivalent to posit that the position be filled, even though the entity that satisfies the EPP (in languages other than English) has neither phonological nor semantic content and is not, therefore, interpretable at the interface levels. Moreover, under these considerations it is not clear what happened to the minimalist statement that movement projects the specifier in the first place: since there is nothing to move there, it is not clear how Spec,TP is created.

In the previous subsection I had discussed the expletive constructions in Capeverdean. Let me recall that the language presents a challenging association of facts:

\begin{enumerate}
  \item i) no overt verbal agreement morphology in any tense;
  \item ii) no overt expletives;
  \item iii) no referential \textit{pro} licensed by Agr (as shown in the first two sections of this chapter, which addressed referential subjects and the pro-drop parameter).
\end{enumerate}
This has led me to the proposal that the way out of this puzzle must be indeed entwined with the canonical subject position, Spec,TP.

### 5.4 Conclusion: Capeverdean status

In this chapter null subjects in Capeverdean finite clauses have been addressed.

First of all, null referential subjects and the conditions for their licensing: according to the empirical evidence and to the diagnostics available in the literature, there seems to be no reason to posit that the empty category *pro* as licensed by agreement exists in the language.

Secondly, expletive constructions, as well as other environments where no subject needs be in Spec,TP (and which had been treated before as performing instances of null expletive subjects, or expletive *pro*): in the former case, there is no subject at all in the sentence; in the latter the subject is allowed to stay somewhere else (such as unaccusatives with indefinite internal arguments or every context where the subject is a clitic, which surfaces on T, leaving no material to surface on Spec,TP).

One important question on the side: in which measure is the inexistence of overt expletives connected with the label pro-drop? I have tried to provide evidence in favor of the separation between both types of “null subjects”.

In section 5.1, the relevant Capeverdean contexts with respect to the null referential subject hypothesis have been presented, both in matrix (5.1.1) and embedded clauses (5.1.2). The conclusion has been that the language allows no null subjects in matrix finite clauses, neither in some embedded clauses where the subject could easily be inferred from the context.

In section 5.2 I have summarized the theoretical evolution on the topic of the null subject parameter, both on the proposals for what is involved in a positive value for it, like the licensing role of rich verbal agreement morphology, and the possible relation to other types of parameters.

In section 5.3 expletive contexts have been addressed: crucial data have been provided (5.3.1) and the EPP has been discussed (5.3.2).
Summing up, in Capeverdean there is:

i) no overt verbal agreement morphology;

ii) no empirical reliable testimony of referential pro in matrix clauses;

iii) no reason to posit an expletive pro;

iv) independent evidence to allow for the proposal of an EPP parameter (if EPP does exist at all), since in the language such Extended Projection Principle does not seem to hold (hence, the value of the parameter is negative);

v) a specific type of embedded context where a null subject is obligatory for the co-reference with a higher operator (a wh-expression or a quantifier DP) to obtain – in these cases, thus, this null subject is a bound variable, not a pro licensed by Agr (the latter, as said above, is prohibited).

Therefore, the conclusion of this chapter is twofold:

i) Capeverdean is not a pro-drop language;

ii) the lack of an overt expletive obligatory in the subject position does not entail the pro-drop status of a language.

In the next chapter, we will enter more deeply in the verb domain. Capeverdean embedded predicates will be discussed and two core questions are addressed:

i) since the language shows no specific morphology for non-finiteness, how can we know whether these embedded verbs are typically non-finite, as expected, or rather display a specific form of (non)finiteness available in these complex contexts only?

ii) in which form is the non-finiteness of these embedded verbs related to the Tense information structure (in other words, can these infinitival forms introduce a temporal reference of their own)?

I will try to solve this by the description and analysis of Capeverdean data, and also by searching for clues in other studies of the same sort. This is the case of Gonçalves 1999, on complex predicates in European Portuguese, and of Wurmbrand 2005, to appear, on English.
Chapter six

Embedded non-finite environments

This last chapter is on embedded predicates that are expected to be infinitival forms. This uncertainty in addressing the infinitive label is due to the fact that, in Capeverdean, there is no specific verbal infinitival morphology. Hence, the diagnostics has to be grounded on other information levels. Such as:

i) the syntactic structure of the embedded clause (following clues provided by the functional material that is allowed in it – namely TMAs and negation);

ii) the temporal reference contained in these embedded environments.

These open questions around the visible materials which a non-finite clause may display are described in 6.1: in 6.1.1 the relevant Capeverdean facts are presented; in 6.1.2 some previous generalizations are provided.

In section 6.2 the list of questions will be narrowed, focusing mainly on the availability of an independent tense. In 6.2.1 I will concentrate on developing the tools that will help apply these diagnostics on non-finiteness. Grounded on the evidence that neither morphological markings nor most of the observations on clausal structure are sufficient to identify Capeverdean embedded infinitives – and, hence, the temporal reference involved must be analyzed in greater detail –, two distinct studies on infinitival tense will be summarized. In 6.2.2 I shall recapture the main ideas and present Capeverdean predicate idiosyncrasies.

In section 6.3 a different matter on embedded infinitives will be at issue: the subjects that are allowed/obligatory in some of these contexts. A few questions concerning the null category PRO will be discussed, including the alternative approach on control as a result of movement, in a fashion somewhat (although not totally) akin to raising, as proposed in Hornstein 1999 and subsequent work.
The main conclusion (in 6.4) will be that, coherently with what has been assumed along this dissertation, Capeverdean Spec,TP is only projected when needed. In some embedded non-finite contexts even T is not projected: the embedded structure is limited to VP. In others, where there is in fact an embedded TP, we verify that these can be analysed as corresponding to English gerunds. In either case, the embedded predicate cannot introduce an independent temporal reference.

### 6.1 What is an embedded infinitive?

In Capeverdean, there is no evident morphological sign of non-finiteness. Hence, the diagnostics has to be grounded on two different strategies:

1. the syntactic structure of the embedded clause and
2. the temporal reference contained in these embedded environments.

In this section I shall organize a list (much probably not exhaustive, though) of the materials admitted in different embedded clauses, according to the selectional properties of the matrix predicates. In subsection 6.1.1 some relevant data will be presented. In subsection 6.1.2 the most evident generalizations will be summarized.

#### 6.1.1 Distinct predicate idiosyncrasies

In (323b) we have an environment which in English is known as containing an embedded infinitive (contrast with the English version).

**Raising:**

(323)  

a. *Ta parseba ma ningen ka staba kasa.*

‘It seemed that nobody was home.’

b. *Ningen ka ta parseba sta kasa.*

‘Nobody seemed to be home.’

---

190 In order to test the ill-formedness of this, I have chosen to maintain the negative concord in Capeverdean, so that we can be sure that the source of the ill-formedness is not caused by any other violation in the language.
In Capeverdean the DP is not allowed to raise to the matrix clause. The only version for these environments is with a matrix *parse*-clause with no subject, embedding a finite full clause which is introduced by the complementizer *ma*.

One of the other predicates widely attested as embedding infinitival clauses is ‘promise’, which in English is an instance of subject control. Before proceeding with the details, we may illustrate the correspondent predicates in Capeverdean.

**Purmete ‘promise’**

Much in the fashion of its English and European Portuguese counterparts, *purmete* ‘promise’ may select for an embedded CP or not. Capeverdean *purmete*, though, has some particular properties of its own:

- when the indirect object is overt, the predicate obligatorily selects for a CP;
- this means that the embedded clause is introduced by *ma* and must display an overt subject and the necessary TMA markers for the sentence to be well-formed.

Check the examples in (324):

(324)  

a. *N purmete nha mai ma-N ta studa pa testi.*

I promise my mother COMP I TMA study for exam

‘I promised my mother that I will study for the exams.’

b. *N purmete nha mai ma-N ta studaban pa testi.*

I promise my mother COMP I TMA study-TMA for exam

‘I promised my mother that I would study for the exams.’

c. * *N purmete nha mai ma-N studa pa testi.*

d. * *N purmete nha mai ma ta studa pa testi.*

e. * *N purmete nha mai ma studa pa testi.*

In both (324a) and (324b) the embedded clause has an overt subject (which needs not be co-referential with the matrix subject – we may say *N purmete nha mai ma nha fidju ta*

191 The same sentence in English might be ‘I promised my mother to study for the exams’ and in European Portuguese *Eu prometi à minha mãe estudar para os exames.*
bai undi-l ‘I promised my mother that my son is going to visit her’), and is correctly marked for a posterior tense with respect to the reference time (in a. it is also posterior with respect to the utterance time; in b. it may be anterior or posterior to the utterance time); both elements (the subject and the relevant TMA) are correct, since we have here a full clause (and the language has no referential pro) and ‘promise’ is a prospective verb, hence this particular embedded tense reference must be expressed.

In the c. example, the sentence is ill-formed because of the lack of the relevant TMA marking (ta, which assures the posterior tense), in d. the embedded subject is missing, and in e. both are missing.

Fla ‘say’ / ‘inform’ / ‘tell’ / ‘promise’ / ‘suggest’
If instead of purmete ‘promise’ we had fla ‘tell’, also with an overt indirect object, this prospective reading – much in the sense of ‘promise’ – has some implications in the embedded details: the obligatory complementizer ma, the obligatory subject and, more specifically (since with the first two the ‘say’ reading may also obtain) the correct temporal markings.

(325) a. N fla nha mai ma-N ta studa pa testi.
‘I said my mother that I will study for the exams.’
b. N fla nha mai ma-N ta studaba pa testi.
‘I said my mother that I would study for the exams.’
c. N fla nha mai ma-N studa(-ba) pa testi.
‘I said my mother that I (had) studied for the exams.’
d. * N fla nha mai ma ta studa pa testi.
e. * N fla nha mai ma studa pa testi.
f. * N fla nha mai studa pa testi.

In (325c) we see that, with certain temporal markings in the embedded declarative clause, fla has the ‘say’ / ‘tell’ meaning (and not ‘promise’): studa, given that it is a non-stative verb, can be either marked with a zero morpheme or with -ba, with both implying anterior readings (the second is more clearly built with respect to the reference time). If the embedded studa were marked by sata(+ba), this would imply a progressive (simultaneous) reading. On the other hand, if it were marked by ta or ta+-ba, it would have a prospective reading (much like ‘promise’ – cf. (325a,b)).
This relation of *studaba* with tense markers, and of this combination with the matrix verb meaning, is not changed if the sentence is a question. In (326), where the lack of *ta* precludes the prospective feature, we observe again that the only compatible reading is ‘say’ (with an embedded relative past), not ‘promise’:

(326)  *Bu fla bo mai ma-u studaba pa testi?*

‘Did you tell your mother that you had studied for the exam?’

* ‘Did you promise your mother that you would/will study for the exams?’

Negation, too, has no effect in changing these effects:

(327)  *N fla nha mai ma-N ka studa pa testi.*

‘I said my mother that I didn’t study for the exams.’

Notice that, in fact, the embedded tense must be anterior (although not marked for it), with respect to the matrix PastPer. These temporal relations are dependent, firstly, on the selectional properties of the matrix verb, according to its precise meaning. When it means ‘promise’ a prospective embedded temporal reference is obligatory (hence, PastPer is ruled out). On the other hand, if the *fla* meaning is ‘say’, we might expect that the temporal properties of the embedded verb also take an important role.

We may test this with an embedded verb that denotes a state (although it is not a “stative verb”). Internal properties of *ten* ‘have’ (as a durative quality, not momentary), allow for various compatibilities of this embedded verb marked with a zero morpheme\(^{192}\) with the matrix *fla* ‘say’. This may be related to the fact that the time of ‘have’ might overlap the moment of the ‘say’ event – for this to be true with verbs that do not denote states, the PastPro markers *sata\(^+-ba* would be needed (as we have seen: *N fla nha mai ma-N sata studaba*... ‘I told my mother that I was studying...’). Check this contrast with *ten medu* ‘be afraid’ (328).

(328)  a.  *N ta flaba bo m’e (ka) ta temba medu.*

   1SG TMA tell 2SG COMP-3SG NEG TMA have.TMA fear

   ‘I used to tell you that he used (not) to be / would (not) be afraid.’

\(^{192}\) Recall that most verbs denoting states – as long as they are not “stative verbs” – have a PastPer reading when combined with a zero morpheme.
b. *N fla-u m’-e (ka) ten medu.
   1SG tell-2SG COMP-3SG NEG have.TMA fear
   ‘I told you not to be afraid.’

c. *N fla-u /ta flaba bo m’-e (ka) ten medu.
   1SG tell-2SG /TMA tell 2SG COMP-3SG NEG have fear
   ‘I told you / used to tell you that he is/was (not) afraid.’

Unlike *purmete, fla does not allow for the version without the complementizer, not even when the indirect object is not overt:

(329) a. N purmete studa pa testi. / N purmeteba studa(-ba) pa testi.
   b. *N fla studa pa testi. / *N fla studa(-ba) pa testi.

Also unlike *purmete, fla may select for a different complementizer: the preposition pa – in this fla + pa context, the ‘order’ reading obtains. Contrast (328), above, with (330) – as before, the boldface distinguishes the relevant complementizer and the TMAs allowed:

(330) a. N ta flaba bo pa-u ka *(ta) ten(eba) medu.
   1SG TMA tell 2SG PA-2SG NEG (TMA) have(TMA) fear
   ‘I used to tell you not to be afraid.’
   b. N fla-u pa-u ka *(ta) ten(*eba) medu.
   1SG tell-2SG PA-2SG NEG (TMA) have(TMA) fear
   ‘I told you not to be afraid.’

Let us now check some possible combinations for *purmete when the embedded verb is not introduced by ma (there is no complementizer), in which there is also no indirect object, no embedded subject and no embedded TMAs other than – in certain conditions – the affix -ba. Contrast the following:

---

193 Thus, only in certain circumstances (with certain embedded verbs) can we apply to fla the observation (for instance in Thráinsson 2003:169) about some English bridge verbs, whose that-complements tend to show main clause-like properties. This is true for Capeverdean fla only when it has a ‘say’ meaning and with certain embedded predicates. Otherwise it only shows some main-clause-like properties: obligatory subject, impossible non-finiteness. The main-clause property that is missing in its that-complements is an independent tense.
   1SG promise (*2SG) (*1SG) (*TMA) give-2SG POSS.2SG book until tomorrow
   ‘I promised to give you your book until tomorrow.’

   b. *N purmeteba daba bo (bo) livru ti manhan.
   1SG promise.TMA give.TMA 2SG (POSS.2SG) book until tomorrow
   ‘I promised to give you your book until tomorrow.’

(332)  a. *N purmeteba ka faze (*fazeba) batota na djogu di manhan.
   1SG promise.TMA NEG make (*do.TMA) cheat in game of tomorrow
   ‘I promised not to cheat in the game of tomorrow.’

   b. *N purmeteba ka fazeba batota na djogu di onti.
   1SG promise.TMA NEG do.TMA cheat in game of tomorrow
   ‘I promised not to cheat in the game yesterday.’

Notice that for the embedded verb also negation is allowed. Therefore, I propose that
the embedded clause is a TP, which cannot have an independent tense of its own: the
prospective (relative future) reading obtains without the marker *ta.*

As for *fla,* its complete meaning (‘say’, ‘inform’, ‘tell’, ‘suggest’, ‘promise’, ‘order’,
‘ask’) is defined by the materials that follow. This may be the reason why when it
selects for an embedded infinitive, it is one introduced by *pa,* and not a bare TP
(wheras with *purmete,* which encodes the prospective tense of its complement, this is
possible). With *fla,* since the temporal reference of the embedded infinitive is dependent
on the matrix predicate, a bare defective TP does not offer the necessary support for *fla*
to build any particular meaning. The complementizer *pa* brings this contribution.

When introduced by *pa,* various prospective meanings of *fla*194 are available
(‘tell’, ‘order’, ‘ask’). I suggest, therefore, that in these embedded infinitives the
preposition selects for a defective TP as its complement. The distinction of this TP
regarding the one selected for by the *purmete* clause is that in the case of *pa*
constructions sometimes the embedded subject is obligatory.

---

194 This includes, thus, most of its meanings, except for the ‘inform’/’say’ – which needs the complementizer *ma.*
And here we arrive at an interesting phenomenon: although some informants detect a certain degree of marginality in subjectless embedded clauses introduced by *pa, their rejection is much stronger when these subjectless clauses have one of those reflexive entries with no overt reflexive expression:

(333)  *N fla Pedru p'-e / *pa ben djanta
I tell Pedru COMP-3SG/COMP come dinner
‘I told Pedru to come dinner.’

(334)  a. *N fla Eder p'-e / *pa ka xuxa ropa.
1SG tell Eder COMP-3SG/COMP NEG dirty clothes
‘I told Eder not to dirty his clothes.’

b. *N fla Eder p'-e / *pa ka xuxa.
1SG tell Eder COMP-3SG/COMP NEG dirty
‘I told Eder not to get dirty.’

(335)  Maria fla Djon p'-e / *pa xinta junto d’-el na testi.
Maria tell Djon COMP-3SG/COMP sit near of 3SG in exam
‘Maria told Djon to sit near her during the exam.’

We shall return to these in the last section of this chapter.

6.1.2 Some generalizations

In this subsection I will summarize some properties of different multi-verb constructions.

Before that, however, it is necessary to discard a possible line of reasoning on these contexts. Given that Capeverdean is a Creole language, and that an equally Portuguese-based Creole as São-Tomense displays serial verb constructions, I must justify why these multi-verb constructions cannot be treated as SVCs.

One might get confused at first glance by some common properties, though. As well described in Hagemeijer 2000, São-Tomense serial verb constructions display double Aspect marking.
The fact that Capeverdean constructions may display -ba on two or more verbs (N staba ta odjaba Maria ta badjaba ‘I was seeing Maria dancing’) might lead to the conclusion that these are similar to serialization. In Hagemeijer 2000 there is a list of criteria, extracted from the literature, to identify SVCs:

- two (or more) verbs expressing a single event;
- a single overt subject;
- one tense marker on V1;
- one negation marker on V1;
- only one aspect marker on V1 or on both verbs;
- no subordinate or coordinate conjunctions;
- no discourse pauses.

Some of these criteria might apply also to multi-verb constructions other than SVCs. Hence, we must go further in the distinctive properties, in order to isolate the distinct patterns. One of the crucial distinctions – and since the double Tense/Aspect marking might be the most confusing feature – is that, as the author points out for São-Tomense, with non-serializing verbs this marking on V2 either is obligatory or prohibited. This is the case with aspectual auxiliary constructions, overlapping events or causative verbs, which are not SVCs. With perception verbs, the Aspect marking on V2 is optional but “crucially independent from aspect marking on V1, and still does not induce the iterativity effects”.

In the proposed generalization of São-Tomense SVCs, the author proposes that these are combinations of lexically restricted heads fusing two argument structures into a single one. The syntactic head, V1, is semantically underspecified, with V2 being the semantic head, imposing s-selection restrictions on the syntactic head. On the other hand, the
verbal status of V2 may be weakened by V1, and ultimately it can be recategorized (it might become relabelled as a preposition, for instance, giving way to a VP-PP configuration instead of the current VP-VP).

The distinction between these properties and Capeverdean contexts with two or more verbs could be detailed to exhaustion. Nevertheless, it suffices to focus our attention on the clear reference to more than one event, distinctly provided by each of the verbs involved. This is so even in the contexts where only one subject is available and the marking -ba is available for both verbs. As we have seen in Chapter Two, the -ba on the second verb is not a repetition of the first, but rather it marks a specific reading for the sentence:

(338)  a. Djon kre ser nh’amigu.
       Djon want be my.friend
       ‘Djon wants to be my friend.’

       b. E kreba ser nha mos.
       ‘He wanted to be my boyfriend [and that is why he asked me for a date].’

       c. E kreba serba nha mos. (counterfactual)
       ‘He wanted to be my boyfriend [but he has no chance].’

Moreover, the temporal localization and nature of the embedded predicate, although being dependent on the semantics of the matrix construction (this is what we shall show in the next section), has its own specific requirements (motivated by the Aktionsart of the predicate in question, for instance).

Now we must turn to some generalizations on different verbs that select for an embedded predicate:

- *pidi* ‘ask’
  - similar to *fla* with the ‘ask’ reading;
  - most of the times it must have an indirect object, except for the cases where the *pa*-complement is a passive.
(339)  a. *N pidi nha pai ku nha mai pa* / p’-es *da-m dinheru pa sinema.*
    ‘I asked my mother and father to give me money for the movie.’
    b. *N pidi pa-N ka fladu kel-la.*
    ‘I asked not to be told that.’

The verbs that we have described so far – *purmete, fla* and *pidi* – can, in other circumstances, clearly select Double Object Constructions:

(340)  *Es purmete-m kel-livru nobu di Nodi.*
    ‘They promised me the new Nodi book.’

(341)  *Kel rapas fla-m nubidadis.*
    ‘That/the boy told me the news.’

(342)  *Mininus pidi-m xokolati.*
    ‘The children asked me for chocolate.’

For the three of them we can have the passivization, even when there is an embedded clause instead of a direct object DP. The indirect object raises to the subject position:

(343)  *N purmetedu ma nha fidju ba pasa di anu.*
    ‘I was promised that my son is going to pass the year.’

(344)  a. *N fladu ma Zé ten un rapariga.*
    ‘I was told that Zé has a lover.’
    b. *N fladu pa-m tra karu di la.*
    ‘I was told to take the car away from there.’

(345)  *N pididu pa-m trabadja dia di dimingu.*
    ‘I was asked to come to work on Sunday.’

The facts above suggest that in these cases we might have a ditransitive construction, for which a sentential direct object is allowed. Recall from the last chapter that the
structured proposed in the present work for the Capeverdean DOC is the following (from (264)):

(346) Capeverdean DOC: \( N \, da-s \, el \):

```
                      VP
                        N    V'
                            V'      DP
                                V       DP
                                   |       | el
                                   da     s
```

The structure in (346) is not problematic at all for the sentences in (340)-(342).

As for the cases where in the place of the direct object we have an embedded clause, this projection would appear in the proper configuration, that is, as an adjunct to \( V' \). This would be the case either when the embedded clause is a CP (as the ones introduced by \( ma \), which in turn selects for a full finite TP) or a PP (in which the preposition/complementizer selects for a defective, non-finite TP). The latter, introduced by \( pa \), displays indeed what seems infinitival tense, but allows, despite that, an overt subject (a clitic, since it is co-referent with either the matrix subject or object, and a full DP is ruled out in these bound positions), the negation and the affix -\( ba \).

We shall analyze the relation between this embedded infinitival tense and the matrix tense in the next sections. The conclusion will be that these infinitives are tenseless. Thus, it cannot be argued that this embedded subject is licensed by some Case feature on T (in the fashion of the Null Case proposed in Chomsky & Lasnik 1993 for PRO). The only structural licensing condition available here is the relation with the preposition \( pa \).
We would have:

(347) DOC-like - *Pedru purmete-m ma Djon txiga oxi.*

```
VP
   
Pedru  V'
   
   V'   CP
   
   V  DP  ma Djon...
   |
   purmete -m
```

(348) DOC-like - *N fla Ana p'-e ka skisi di txoma-m:*

```
VP
   
N  V'
   
   V'   PP
   
   V  DP  p'-e ka skisi...
   |
   fla   Ana
```

(349) DOC-like - *N pidi nha mai p'-e da-m dinheru:*

```
VP
   
N  V'
   
   V'   PP
   
   V  DP  p'-e da-m dinheru
   |
   pidi nha mai
```
As for the contexts where purmete does not have an overt indirect object, in which case it selects for a complement with no complementizer, we must have an embedded TP, even though this is defective as well. This is so since we can find the negation ka and the affix -ba (not the subject, though – recall that the subject above could be licensed by the preposition; here it is ruled out – this absent embedded subject may be accounted for by way of positing PRO or by arguing it has been moved to the matrix Spec,VP):

\[(350)\]

\[N \text{ purmeteba ka fazeba batota:}\]

\[
\begin{array}{c}
\text{VP} \\
\text{N} \\
\text{V'} \\
\text{V} \\
\text{TP} \\
\text{purmeteba ka fazeba batota}
\end{array}
\]

Observe some different combinations:

*manda* ‘order’ and *dexa* ‘allow’/’let’
- there is an overt indirect object;
- with *manda* there may be, or not, a complementizer introducing the complement;
- when there is no complementizer, there are further restrictions to the temporal reference of the embedded clause;
- with *dexa* there is never a complementizer;
- when the subject of the embedded verb is a pronominal form, it cliticizes to the matrix verb, when the morpho-phonological conditions allow this – that is, not with the -ba affixed to the matrix verb (check Chapter Four for the final “-ba problem”); hence, we have *Es dexa-m bai*\(^{195}\) / *Es dexaba mi baba*; *Es manda-m bai* / *Es mandaba mi baba*;
- thus, it looks pretty much like an ECM environment;
- the only material allowed in the embedded verb domain is *ka* and -ba in the case of *dexa*; in the case of *manda* with no embedded *pa*, only -ba is allowed.

\(^{195}\) We may have more complicated contexts, like *Mi-N fla nha pai ku nha mai p’-es dexa-m ba odja filmi* ‘I asked my father and my mother for them to let me go see the movie’.
(351)  

   ‘Today it is raining a lot, I let them not go to school.’

b. *N mand’-es ka bai skola.

The meaning ‘I ordered them not to go’ is given with *fla: N fla-s p’-es ka bai.

In (352) we have three examples with manda.

(352)  

a. Si pursora sabeba ma-N ta txigaba, e ta mandaba mininus xintaba.
   ‘If the teacher knew that I would arrive, she would order the children to sit.’

b. *Kantu N txiga pursora sata mandaba mininus xinta(*-ba).

b’. Pursora sata mandaba mininus xinta(*-ba) kantu N txiga.
   ‘When I arrived the teacher was ordering the children to sit.’

c. Pursora sata mandaba mininus p’-es xinta(-ba) oki N txiga.
   ‘The teacher was ordering the children to sit at the moment I arrived.’

The reading in (352c) is only obtained in this configuration, not in (352b’): in this one, my arrival is a moment coincident with the ordering on-going event; in (352c) the on-going event is anterior to my arrival, and the teacher wants that my arrival is coincident with the children sitting (recall that this a distinction sustained by the difference between *kantu – past or habitual – and *oki – future and episodic).

Besides these verbs, we also have modal auxiliaries, which behave as expected:

\[\text{pode} \begin{array}{l}
\text{‘be able’ / ‘be allowed’ (both modal readings)}^{196}
\end{array}\]

- \[\text{ka and -ba are allowed for the embedded verb, depending on the desired reading:}\]

(353)  

\[\text{Bu podeba ka ba ku mi.}\]
   ‘You had the permission not to go with me.’

(354)  

a. \[\text{Bu ka debe fla kasi pa bo mos.}\]
   2sg neg mod say/tell lie prep your boy
   ‘You must not lie to your boyfriend.’

\[\text{As a side note, there is no clitic climbing in the language:}\]

(i)  

a. *\text{Bu pode-l leba.}\]

b. \text{Bu pode leba-l.}\]
b. * Bu debe ka fla kasi pa bo mos.
   2SG MOD NEG say/tell lie prep your boy

(355) a. Es mininu li ka pode bebe lete.
   DEM boy LOC NEG can drink milk
   ‘This boy cannot drink milk.’

b. Pode txobe, pode ka txobe, mi go!
   MOD rain, MOD NEG rain, 1SG now
   ‘It may rain, it may not rain, I don’t care!’

Finally, we have a different type of embedding predicates, the ones that allow for an embedded *ta*. This, however, does not seem to preclude the infinitival status of the clause, since the reading is, as we have seen earlier, the equivalent to a gerund form.

   1SG TMA see-CLITIC3SG TMA dance
   ‘I see him/her dancing.’

b. ??N ta odja el ta badja.
   1SG TMA see 3SG TMA dance
   ‘I see he/she dances.’

c. N ta odjaba el ta badja(ba).
   1SG TMA see.TMA 3SG TMA dance
   ‘I was seeing him/her working.’

(357) a. N ka gosta di obi nha fidju ta txora.
   ‘I don’t like to hear my daughter crying.’

b. E ta lembra di odja si mai ta badja.
   ‘He remembers seeing his mother dancing.’

Again, we have what might be considered an ECM construction. On the other hand, we might consider that the *ta*-clause – like some *ing*-clauses in English or some prepositional infinitives in European Portuguese – predicates over the object.
Now that some generalizations have been made, and the hypothesis of dealing with multi-verb constructions as if they were serial verb constructions (akin to what has been defended for other languages, including São Tomense) has been discarded, it is necessary to discuss the status of the embedded clauses with respect to the temporal reference they can assume.

This is crucial for the present work since, as has been shown, Capeverdean does not display any specific infinitival morphology.

The second diagnostics proposed for the classification of these embedded constructions as infinitives has been the material allowed in the embedded functional domain. This does not seem to help clarify this matter, though. We have cases in which the negation, or the subject, or *ta* or *-ba* are allowed; in some cases some combinations among them are possible.

In the end of the list of diagnostics, we have a more complicated, less visible one: the temporal reference of the embedded clause. Following the lines of reasoning in the works of Gonçalves 1999 and Wurmbrand 2005, the embedded infinitivals are tenseless: their temporal reference is derived from the matrix tense. As a last resort, so, I will apply this to Capeverdean and show (we have had some strong clues on this) that these embedded clauses are indeed infinitives. In other words, even considering that there are different types of infinitives, the temporal reference they introduce is dependent on the matrix clause. In this sense, they are tenseless.

### 6.2 Infinitives and tense

In this section I shall try to identify which, among the relevant Capeverdean predicate embedded clauses, are infinitival contexts. This will be achieved with recourse to the second diagnostics that seem to apply and which can be subsumed to the following question: can these introduce an independent tense?

In subsection 6.2.1 the lines in Wurmbrand (2005, to appear), on English infinitives, and in Gonçalves 1999, on European Portuguese complex predicates, will be summarized, in order to gather some crucial concepts that will inspire an approach to Capeverdean. Both these authors contend that infinitives are tenseless. I shall apply to Capeverdean a reverse line of reasoning: whenever the embedded temporal reference cannot be independent from the matrix tense, and when this condition cumulates with
the lack of specific tense requirements on the part of the matrix verb\textsuperscript{197}, the embedded clause in question is non-finite; whenever an embedded independent temporal reference is allowed, the embedded clause is finite\textsuperscript{198}.

This will be the main goal in 6.2.2, where Capeverdean facts are analyzed in this light, using, among other empirical details, a new problem concerning the affix -$ba$.

6.2.1 Two previous approaches: infinitives are tenseless

According to both Wurmbrand (2005, to appear) on English, and Gonçalves 1999 on European Portuguese complex predicates, embedded infinitives cannot introduce an independent reference for tense: in other words, they are truly tenseless, against some views available in the literature.

The main line in Wurmbrand (to appear) goes against the existence of an infinitival tense. I will summarize her descriptions and arguments, since they are crucial for what I will propose regarding Capeverdean.

Since Stowell 1982 some authors have considered two types of infinitival complements: tensed and tenseless infinitives, whose presence/absence correlates with different syntactic structures or properties. Concentrating on future irrealis infinitives, like ‘Leo decided to go to the party tomorrow’, which have been uniformly considered to be ‘tensed’ infinitives\textsuperscript{199}, the author provides further evidence for the lack of tense in these, thus agreeing with a common assumption in many semantic works on infinitives and going against the mentioned accounts that involve infinitival tense.

\textsuperscript{197} Take for instance purmete ‘promise’: even when it has an overt indirect object, thus selecting for a $ma$-complement, it maintains its temporal selectional properties — a prospective embedded tense is required. This requirement does not mean, however, that the embedded clause is non-finite.

\textsuperscript{198} For the approaches that consider a possible distinct tense in embedded infinitives, the first part of this proposal might not be controversial: we only have to assume that the tenseless infinitives are a subset of all infinitives and, thus, I am referring to this tenseless subset (which is about infinitives anyway). As for the second part — whenever there is an independent tense, what we have are finite clauses — the problem could be more complex. However, as we shall see, the only cases in which this happens are some of those introduced by the complementizer $ma$, non-controversially assumed to surface on C and selecting a full clause, with obligatory referential subject and the temporal morphology required by the lower predicate in question. Recall, from what has just been said about the inherent temporal requirements of, at least, purmete ‘promise’, that not all the embedded clauses introduced by $ma$ allow for an independent tense.

\textsuperscript{199} Wurmbrand (to appear) notes that there is disagreement about others, such as factives or $try$ constructions: the first are considered to be tenseless by Pesetsky & Torrego and tensed by Landau and herself; the latter are considered as tensed by Pesetsky & Torrego and tenseless by Landau and herself.
Following Abush 1985, among others, Wurmbrand (to appear) considers future as composed by two parts: tense (present tense) plus the abstract modal *woll* (yielding posterity). These parts are morphologically spelled out as *will*. One argument in favor of the composite nature of *will* involves certain sequence of tense (SOT) effects; the other argument is supported by the indexical or absolute nature of the English future.

The reasoning under the latter goes as follows: English present time – *Pres* – is indexical/absolute (it must contain the utterance time; see Enç 1987, Abusch 1988, a.o.), as opposed to languages whose *Pres* is defined as relative (in the latter, for a sentence like ‘Leo found out that Mary is pregnant’ there is a possible reading where the time of pregnancy overlaps the matrix time – the founding out time – but does not reach the utterance time); thus, in English *will* constructions, which likewise must be evaluated with respect to the utterance time (in ‘Leo found out that Mary will be pregnant’ the time of pregnancy must be after the matrix time and after the utterance time), “the absolute nature of the sentence follows straightforwardly if it is assumed that *will* decomposes (syntactically/semantically) into two parts, a future modal and an indexical/absolute *Pres* tense.” (Wurmbrand, to appear:2)

As for the SOT phenomenon (see Dowty 1982, Abusch 1988, Ogihara 1996, a.o.), it refers to certain contexts where a morphologically realized tense is semantically vacuous. For instance in ‘Leo found that Mary was pregnant’, where the embedded clause can receive a non-past interpretation, the pregnancy time is not in the past with respect to the matrix – the founding out – time but rather overlaps it. Wurmbrand follows Ogihara 1996 proposal that a tense may delete at LF if it is in the scope of another tense with the same value (the embedded *Past* is in the scope of another *Past*). Semantically, and following Heim 1994, the deleted tense variable is interpreted as relative *now* (the embedded time is simultaneous to the matrix time). This is crucial in considering the *Pres* component of English *will* in that an embedded *Pres* may be interpreted as a zero time – a relative *now* – with respect to a matrix *will* construction, such as in ‘John will see the unicorn that is walking’ (where *is walking* is not simultaneous to the utterance time, but simultaneous to the matrix time). Thus, there must be a *Pres* tense encoded in the *will* construction.

Wurmbrand departs from these semantic observations on future in order to show that, unlike English finite constructions (where future is absolute), English infinitival future is relative (the embedded event must occur after the matrix event, but it may occur before the utterance time). Wurmbrand suggests that this is due to the fact that
Pres tense assures that the future is absolute and the lack of tense results in relative future. We can see this in the following examples: control (358b) and ECM (359b).

(358) a. Leo decided a week ago that he will go to the party (*yesterday).  
    b. Leo decided a week ago to go to the party yesterday.

(359) a. According to a report I read last week, it was expected that the bridge will collapse (*yesterday).  
    b. According to a report I read last week, the bridge was expected to collapse yesterday.

This allows us to conclude that infinitival future only has the *woll* part, not the Pres part\textsuperscript{200}. The presence of *woll* (recall that this contributes the meaning posterity) in infinitival future is also attested by some examples showing that it is *woll* that licenses eventive predicates in the future. We can see this by the following contrast:

(360) a. Leo sings in the shower (*right now)\textsuperscript{201} PRES:*  
    b. Leo sang in the shower right then. PAST: ✓  
    c. Leo will sing in the shower right then. FUT: ✓

Involving *woll*, and not Pres, infinitival future is compatible with an eventive predicate (we may try adding a tomorrow in the end of (361)):

(361) Leo decided to sing in the shower. (*woll)

As for simultaneous infinitives, certain constructions prohibit eventive predicates (Wurmbrand argues they lack both tense and *woll*). This is the case in (362):

(362) a. Lina believes Leo to sing in shower (*right now).  
    b. Lina believes Leo to be singing in the shower right now.  
    c. Leo claims to play the Marsellaise (*right now).

\textsuperscript{200} Wurmbrand also argues against the views that oppose control infinitives (these contend that, these clauses being tensed, they allow eventive predicates) to ECM/raising infinitives (as tenseless, they do not allow eventive predicates); the author proposes that this is empirically incorrect, since both constructions allow and prohibit eventive predicates, according to the contexts.

\textsuperscript{201} This sentence is OK with the habitual reading, but not the eventive reading.
However, eventive predicates and simultaneity are compatible in infinitives with implicative predicates such as ‘manage’, aspe ctual predicates such as ‘begin’ (whether control or raising) and verbs like ‘try’. Being incompatible with a future reading, they do not involve *woll:

(363) a. Leo managed to sing in the shower (*tomorrow).
    b. Leo began to sing in the shower (*tomorrow).
    c. The tower began to fall over (*tomorrow).
    d. Leo tried to sing in the shower (*tomorrow).

Does this mean that these, besides not involving *woll, involve tense (by contrast with the ones in (362))? Wurmbrand’s answer is no, and she proposes for the type of infinitives in (363) that they are typical predicates triggering restructuring or clause union in different languages (such as in Rizzi 1982). There is a transparency effect associated to restructuring infinitives, since they lack, to different degrees, clausal functional projections (they display a truncated structure). In other terms, lacking the functional structure above $\upsilon P$ (the embedded $\upsilon P$ is the complement of the matrix V), these embedded structures do not project any $woll P$. There is only one TP in the sentence, the matrix TP, and thus the embedded verb receives a simultaneous interpretation.

Furthermore, the event variable of an embedded eventive predicate can be bound by the matrix tense. Like we can observe in (364):

(364) a. Leo seems to sing in the shower (*right now).
    b. Leo seemed to sing in the shower (right then).

In (364a) the matrix tense is Pres, hence there is no eventive licenser involved and the embedded eventive predicate with *seem* is prohibited; in (364b), where the matrix tense is Past, an embedded eventive predicate with *seem* is allowed. This contrast shows two important things:

i) this sensitivity to the matrix tense makes it hard to accept that the embedded infinitive also involves tense;
ii) the crucial distinction must not be in the particular selectional properties of the type of matrix predicates, since in this view the contrast in (364) would be hard to explain as well.

Finally, the author points out that a contrast like the one in (365) may be accounted for if there is no embedded TP:

(365)  

a. There seems to be a unicorn in the garden.

b. * There seems a unicorn to be in the garden.

The other way of explaining (365b) is in the fashion of Bošković 2002: there is no EPP triggering movement of the associate of there, hence movement is disallowed.

The above summary of Wurmbrand (to appear) intends to bring to the present work some of her arguments in favor of infinitivals as being tenseless.

As for Gonçalves 1999, the author argues in favor of the more specific tenseless property of infinitives involved in complex verbal predicates\(^{202}\) in European Portuguese. The author presents a distinction between syntactic complex predicates and lexical complex predicates, and she also points out the fact that some categories other than V can enter this type of construction where a complex head obtains. Concentrating on verbs, the ones that trigger complex predicate formation — since these are of the restructuring type or of the type the author calls fazer-Inf (‘make’-Inf)\(^{203}\) — are always a subset of control (querer ‘want’, but not decidir ‘decide’, for instance), raising (not parecer ‘seem’, but raising modals like poder ‘can’) and causative/perceptive verbs\(^{204}\).

\(^{202}\) A complex predicate is “any inflectional domain that contains two distinct morphemes, each of which selects at least one phrasal argument in its theta-grid.” (Baker 1996:338, quoted in Gonçalves 1999:37)

\(^{203}\) According to classical analyses on the distinction between these two types, this is determined by the internal properties of the matrix verb. However, the class of the matrix verb, although being a necessary condition, is not sufficient to determine restructuring: in each class there are verbs that trigger restructuring and others that do not.

\(^{204}\) The author posits the following condition for the formation of a restructuring complex predicate: “The complex predicate formation is only possible if the behavior of the infinitive is temporally dependent on the temporal specifications of the higher domain […] because it may not contain temporal information in conflict with the above mentioned specifications.” (Gonçalves 1999:243, my translation) She points that there are, however, some verbs (like ousar ‘dare’) that cannot introduce a temporal reference of their own and yet they do not allow complex predicate formation either.
We shall see that the syntactic analysis explored in Gonçalves 1999 for European Portuguese restructuring verbs cannot hold in Capeverdean, since, in order to account for the lack of obligatory adjacency (as opposed to what happens for instance in Catalan), this analysis resorts to the movement of the matrix V to a functional head (whereas the embedded verb moves to the higher V)\textsuperscript{205}. The fact that the two verbal units surface on distinct and non-adjacent heads (and not on the same syntactic position) allows for some material between them, which is borne out by EP data.

The verbs that trigger restructuring may have TP complements (this is the case with control and raising verbs) or VP complements (this is the case with temporal and aspectual verbs). The heads C and AgrS are never projected above the embedded VP (against what had been proposed in the pre-B\textsc{arriers} era). Gonçalves 1999 (as does Wurmbrand 2005, to appear) states that the infinitival complement cannot introduce a temporal reference of its own. The author also contends that the fact of there being no embedded T (with temporal and aspectual verbs), or the fact that T – when projected (with control and raising verbs) – is defective in its V-features, triggers the movement of the embedded verb to the higher V; here, the embedded verb can check its T-features in a relation with the matrix T.\textsuperscript{206}

For the contexts where the verb does not trigger restructuring, the author proposes that the embedded T is active, being able to check the features of the embedded verb, which precludes the complex predicate formation. This shows that the process of restructuring is not optional: the complex predicates are not derived from non-complex predicates; they are simply distinct phenomena.

This process of complex predicate formation with restructuring verbs is similar to the one with the \textit{fazer}-Inf type of verbs. Although there are some differences between the two types with respect to the passive formation and also to the structural composition of the complement (the infinitival complement is more defective with \textit{fazer}-Inf than with the restructuring verbs – in the first, AgrS, T and AgrO are never

---

\textsuperscript{205} This does not destroy the complex head, only allows for each of the units to be spelled-out in different positions. A different possibility mentioned in Gonçalves 1999 would be to consider that the complex predicate is only formed at LF (after Spell-Out, which excludes any consequences at the PF component), but the author’s choice falls on the other one.

\textsuperscript{206} Again, it is not clear which features trigger V movement: the V-features of T or the T-features of V.
projected), they behave in a similar fashion with respect to clitic climbing, long distance movement, clausal negation and non-obligatory adjacency between the verbs (adverbs like *bem* and quantifiers may appear between the two verbal units\textsuperscript{207}.

On the other hand, there are crucial differences between the *fazer*-Inf constructions, on one side, and ECM and inflected infinitival constructions, on the other. Only the first trigger complex predicate formation, also allowing for an inversion in the order of the embedded arguments (this is a productive construction in Romance languages)\textsuperscript{208}. Causative and perceptive verbs also show some differences between them. The author then proposes a scale according to which the *fazer*-Inf construction selects a defective infinitival complement (a Caus projection), followed by the ECM construction, whose infinitival complement is headed by a defective T, and finally the less defective complement of all these three is the one in inflected infinitive constructions, headed by AgrS.

As for the perceptive verbs, there are important implications induced by the properties of the embedded verb: the complex predicate formation is only possible when this one is intransitive. The fact that complex predicates with perceptive verbs are prohibited when the embedded verb is transitive is accounted for in Gonçalves 1999 with thematic assignment incompatibilities: perceptive verbs are responsible for the assignment of the embedded subject theta-role, and the specific theta-role it assigns to it is only possible with mono-argumental embedded verbs. This seems to be confirmed by the possible embedded transitives in complex predicates with causative matrix verbs.

The author concludes that the motivation for complex predicate formation is independent of the class of the matrix verb, since the movement of the embedded verb is due to the defective properties of the infinitival complement. In other words, with the same matrix verb we may have different structures, depending on the embedded syntactic requirements.

In this subsection I have summarized two relevant studies on embedded infinitives – Wurmbrand (2005, to appear) on English, and Gonçalves 1999, on European Portuguese. In the next subsection I will bring the Capeverdean data at issue and argue that the infinitival status of the embedded clauses under discussion follows straightforwardly from the combined informations we have so far.

\textsuperscript{207} These transparency effects are the diagnostics proposed in Rizzi 1982.

\textsuperscript{208} As for ECM, there is an obligatory SVO order in the infinitival complement – just as in English.
6.2.2 Capeverdean possible combinations

From section 6.1 we have some crucial generalizations about certain Capeverdean predicates which select for embedded clauses. Their counterparts in other languages have been described as having infinitival complements. We may combine information on the materials allowed in these embedded clauses and the determination about the non-independence of their temporal reference and check whether these are finite or not.

Consider the table in (366).

(366) Various matrix verbs and their selectional properties.

<table>
<thead>
<tr>
<th>matrix predicates</th>
<th>embedded functional material</th>
<th>independent embedded tense</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ma\textsuperscript{209}</td>
<td>pa\textsuperscript{210}</td>
</tr>
<tr>
<td>modal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pode ‘can’</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>debe ‘must’</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>modal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>manda ‘order’</td>
<td>*</td>
<td>⊥</td>
</tr>
<tr>
<td>daxa ‘allow’</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>modal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>obi ‘hear’</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>cause/allowance / perception (ECM)\textsuperscript{211}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>purmæte ‘promise’</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>kre ‘want’</td>
<td>*</td>
<td>⊥</td>
</tr>
<tr>
<td>control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fla ‘tell’</td>
<td>*</td>
<td>⊥</td>
</tr>
</tbody>
</table>

First of all, notice that there are selectional differences among verbs of the same category, for instance between pode ‘can’ and debe ‘must’ (the former may select for an embedded negation, which has scope over the lower clause, whereas the latter may not). These differences obviously depend on semantic specificities of the verbs in question.

\textsuperscript{209} Selecting for a ma-clause means to have an embedded obligatory subject.

\textsuperscript{210} Allowing for pa also means allowing for the embedded subject; although preferably overt, it is only obligatory in certain reflexive contexts.

\textsuperscript{211} The reference to Exceptional Case Marking here is only because these predicates select for an obligatory DP object; under certain views, this DP, being the subject of the embedded clause, is assumed as being raised to the matrix object position, in order to check Case (since this is unavailable from the embedded non-finite T).

\textsuperscript{212} Manda and kre only allow for the embedded negation in the case they also project a pa-clause.
At a first glance, two facts are exceptionally evident:

i) -ba is allowed in all types of embedded clauses;

ii) besides the contexts where the clause is introduced by ma (hence, it is a full clause, headed by C and displaying a finite Tense), ta is only allowed in the clauses selected by perceptive verbs.

This lower ta does not preclude a non-finite reading. The true difference between these contexts and the others is that in the latter there is a simultaneity reading for the embedded event with respect to the matrix verb (in any case, this is a dependent tense). In other words, this seems to correspond to the English gerund in its present participle constructions; these constructions predicate over the object of the perceptive verb, and in European Portuguese they are built by a + infinitive (prepositional infinitives), like in (367)\textsuperscript{213}:

(367) a. O Zé ouviu o Paulo a cantar.
    b. Zé heard Paulo singing.

Just as in (367), in Capeverdean N odja Maria ta badja ‘I saw Maria dancing’ what we have is a perception verb and a type of gerund construction – with a progressive property – as a complement to the matrix object. We can analyze this as a secondary predicate, although it is not important for the time being. The most relevant point here is that, in these type of environments with a higher perceptive verb (traditionally viewed as ECM), over whose object one ta construction predicates, this ta construction must have a simultaneous interpretation with respect to the perceptive verb.

To obtain a distinct gerund reading, such as in certain adjunct clauses, the Capeverdean construction requires a different marking: na. Check the following:

(368) E kori mundu na konta kasi.
    3sg run world PREP tell lie
    ‘He went on and on telling lies.’

\textsuperscript{213} See Lobo 2003 for more details on English and European Portuguese gerund forms.
Only to illustrate other possible contexts where an infinival clause may be selected, we have the following (notice that all of them pass the proposed test for nonfiniteness: they cannot introduce an independent tense – their temporal reference is always dependent on the composition of the matrix tense + temporal specifications of the verbs involved):

Selected by an aspectual verb:

(369)  *N ten ki panha ropa antis di kumesa ta txobe.*

1sg have rel fetch clothes before of start TMA rain

‘I have to fetch the clothes before it starts raining.’

Selected by an adverb:

(370)  *Antis di txobe ben un grande vendaval.*

‘Before raining there comes strong wind.’

Extraposition also selects for infinitives:\^{214}:

(371)  *E faxi pa bu prendi badja tornu.*

BE easy for 2SG learn dance tornu

‘It is easy for you to learn how to dance tornu.’

And still some more contexts headed by *pa*:

(372)  *Era pa Kadi traze-m el.*

‘It was for Kadi to bring him [the baby] to me.’

\^{214} But not clefts (in which the embedded clause is headed by *ki*):

(i)  *E di dinheru kiN sata mesti.*

‘It is money that I need.’

Neither adjectival predicates:

(ii)  *E kapas ki txuba ta ben.*

‘It is likely that rain comes.’

---

316
(373)  a. N pidi nha xefe pa N sai mas sedu. (the pa-clause is a sentential direct object)  
   ‘I asked my boss to leave early.’  
   b. N tilifona nha xefe pa N konta-l novidadis. (the pa-clause is an adjunct)  
   ‘I called my boss to tell him the news.’

In (373b), pa (just as the English ‘to’ and the European Portuguese para in identical contexts) clearly has a stronger purpose meaning than in (373a).

Some more verbs select for an infinitival complement. Contrast the following strategies (as a side note, meste needs no ta to mark episodic present, only habitual or future):

(374)  a. N fla-u pa bu ka ten medu.  
   ‘I told you not to be afraid.’  
   b. N fla-u ma bu ka meste ten medu.  
   ‘I told you that you need not be afraid.’  
   c. N flaba bo ma bu ka meste ten medu.  
   ‘I had told you that you need not be afraid.’

As for those generic contexts which are traditionally identified with non-obligatory control (hence, a generic PRO is said to be in the subject matrix position), they are built in Capeverdean by way of a different strategy: the impersonal passive.

(375)  a. Ka pode fumadu na kel sala-li.  (Capeverdean)  
   a’. Não se pode fumar nesta sala.  (European Portuguese)  
   ‘To smoke is forbidden in this room.’  
   b. Ka ta fumadu na kel sala-li.  
   b’. Não se fuma nesta sala.  
   ‘Nobody smokes in this room.’

The other hypothesis is a generic 2PL clitic.

(376)  a. Li pode nadadu.  
   b. Li nhu pode nada.  
   ‘Here you can swim.’
These descriptions and analyses of Capeverdean infinitives lead us to some questions:

- is there any PRO in the language?
- are these contexts really akin to the ones traditionally viewed as control?

This is what I will address in the next and last section.

6.3 Is there a PRO in Capeverdean?

The question in the title is grounded on two reasons:

i) given what we know regarding the functional structure of the language, namely that the Spec,TP position is only projected when needed, is there any sense in resorting to a much problematic null category only to fill in this position?

ii) given what we have just observed about the selectional properties of predicates with respect to their infinitival complements, are we actually sure that such a position is needed in the cases where no overt subject is obligatory?

Before addressing these particular questions on Capeverdean, it is necessary to discuss PRO as such problematic a category. This is will be at stake in 6.3.1, reviewing some recent arguments in the literature. In 6.3.2 I shall bring into the discussion a very specific Capeverdean problem: in the contexts where the embedded verb is a reflexive entry with no reflexive expression (the ones studied in 4.2), the subject of the infinitive is obligatory (in the other contexts, although it is preferred, its absence is allowed).

6.3.1 PRO or a trace of DP-movement?

In this section I will address the question on how subjects are licensed in some particular non-finite contexts. I am referring to those contexts that, under a Government and Binding (GB) approach, have been labeled as Control: in the subject position of an embedded non-finite clause, which is a caseless position since T, when projected, is non-finite, what we have is a null category PRO. This is a bizarre category which is simultaneously [+pronominal] and [+anaphoric], and that must be controlled by (or, in
other words, whose reference must be identified with) a DP in the matrix clause; in
certain cases (with certain predicates and under certain circumstances) this controller is
the subject of the matrix clause, in others it is the object.

In the last decade or so, Control has become a “hot topic” (expression borrowed
from Boeckx & Hornstein 2004). Generativists know that the GB considerations and
proposals concerning this special type of subjects are far from satisfactory. But this
more lively and recent debate seems to be due to some specific proposals, like in Martin
1996 and in particular in Hornstein 1999, which have raised a controversial hypothesis:
control structures are the product of syntactic operations similar to Raising; that is, they
are the result of A-movement. This is only possible if one dispenses with the theta-
criterion, thus allowing movement of the DP to a theta-position. Therefore, for an
English sentence like ‘John hopes to leave’, we would have the following facts:

\[
[\text{IP} \text{John} \ [\text{VP} \text{John} \ [\text{hopes} \ [\text{VP} \text{John} \ [\text{leave}]]]]] \quad (\text{Hornstein 1999:79})
\]

‘John’ merges first with ‘leave’ (external merge) and is assigned the theta-role from the
verb; it “raises” to the embedded Spec,IP to check the D-feature of I and, since this is
not a case position, it “raises” again to the matrix Spec,VP, where it checks the external
theta-feature of ‘hope’, assuming this theta-role (‘John’ is now both the ‘hoper’ and the
‘leaver’); finally it “raises” to the matrix Spec,IP, where it checks Nominative Case.
This corresponds to a control analysis in the following terms: ‘John’ in the lowest
Spec,IP is PRO, and ‘John’ in the highest Spec,IP is the antecedent, or the controller.

Notice that, although some critical positions have accused this proposal of trying to
collapse Control and Raising (Landau 2003:471), there has been a later clarification
(Boeckx & Hornstein 2004) on this: “the two instances of movement cannot be
collapsed. There is an important difference between the two. In control, but not raising,
the moving element remerges into a θ-position prior to reaching Spec,TP.” (Boeckx &
Hornstein 2004:432, italics by the authors) These authors add that “The principal MTC
[Movement Theory of Control] strategy is to explain all differences between raising and
control in terms of this one difference.”\textsuperscript{215}

\textsuperscript{215} This controversy concerning PRO and the MTC has, at least, the following “episodes”: Hornstein
The MTC approach seems to fit well on the so-called late minimalist concerns, namely the search for a more elegant and simple theory, for it eliminates some concepts and explanatory strategies that had been created to account for PRO alone. In the literature, “late minimalism” refers to Chomsky 1995 and afterwards, while “early minimalism” refers to papers from the early nineties, like Chomsky 1993. For instance in Chomsky & Lasnik 1993 those elegance concerns were surely not present among the authors’ priorities, since one more stipulation is proposed to solve some of the problems concerning PRO idiosyncrasies: the Null Case, a special Case that would be due to contexts in which T, although non-finite, is tensed (it brings some tense information to the sentence – we shall see this below). This has led to the amusing observation that PRO is so idiosyncratic that it gets more attention than ‘John’ and ‘Mary’, since it has a Case of its own, let alone its own theorem216.

If one assumes the MTC as a possible way out of (at least) some of the puzzles concerning the idiosyncratic PRO, there is still one question raised by Capeverdean facts. Given what I have been arguing on the Spec,TP position (or any functional position) as being projected only when needed, what can we say regarding the subject position of embedded infinitival clauses?

In other words, Capeverdean facts (and the line of argumentation that I have been developing along the present work) lead us to the following questions:

- why do we need the embedded Spec,IP position in the first place?
- either to be occupied by PRO, or to be a landing site for the embedded DP on its way to a case position?
- either way, what role does this position play in the whole picture?

Now suppose that, for the English sentence in (377), only T is projected: it is defective (non-finite and, thus, tenseless) but we know it is there, and it is where the word ‘to’ surfaces. ‘John’ raises from the embedded Spec,VP to the matrix Spec,VP, and then everything else predicted by Hornstein 1999 applies.

Next subsection will add some empirical information that might help illuminate this.

216 In Chomsky 1981 – thus, when government was a core notion –, the PRO Theorem tried to account for the fact that it is simultaneously a pronoun and an anaphor; roughly, the theorem predicted that its position should be ungoverned; this means that PRO is in complementary distribution with overt NPs.
6.3.2 Reflexive contexts revisited

In Capeverdean contexts with an embedded infinitive headed by pa the subject is obligatorily overt if the embedded predicate is an unaccusative or a reflexive entry with no overt reflexive, such as the cases that have been described and analyzed in 4.2.

(378) a. *N fla Djon p'-e / pa ben djanta. (embedded unergative)
    1SG tell Djon to 3SG / to come dinner
    ‘I told Djon to come to dinner.’

b. *N fla Djon p'-e /* pa ben li. (embedded unaccusative)
    1SG tell Djon to 3SG / to come LOC
    ‘I told Djon to come here.’

c. *N fla Djon p'-e / pa laba mo. (embedded transitive)
    1SG tell Djon to 3SG / to wash hand
    ‘I told Djon to wash his hands.’

d. *N fla Djon p'-e laba. / * N fla Djon pa laba. (embedded reflexive)
    ‘I told Djon to wash (himself).’

In (378a,c) the embedded subject can be overt or not; although some informants clearly prefer the version which has it overt, many others do not reject the “silent” subject version; moreover, even the informants who prefer an overt subject in these unergative (the main verb in (378a) is djanta ‘dinner’, and ben ‘come’ is an auxiliary) or transitive entries do not repudiate here the “silent” version so consistently as they – all – do with the “silent” subject in “reflexive” contexts.

Recall from 4.2 that these “reflexive” entries are to be analyzed as constructions where the DP has been generated as internal argument, just as in passives. This is coherent with the proposal in 4.2.2 that the result of these actions falls on the DP subject; cumulatively, there is the presupposition of some action of the DP for the event result to obtain, either on purpose, which corresponds more to a get-reading, or accidentally/involuntarily, which corresponds more to a become-reading (mainly in verbs of the group ii), in (225)). This allows us to propose that these are both instances of passive constructions. This merge of two theta-roles is what leads us to call these “reflexives”, despite the fact that no SELF anaphor is realized.
Now we have some more information that may help illuminate both these constructions and control. One possible explanation for the contrast in (378c) would be by way of assuming Hornstein’s MTC proposal (control as movement). This goes as follows.

Consider that in (378d):

i) the DP Djon, the subject of the embedded infinitival clause, has been generated as the complement of laba ‘wash’ (this is, not controversially, what happens to the DP subject in (378b), with an unaccusative);

ii) as proposed in 4.2, the DP in (378d) has moved from the embedded object position to the embedded Spec,VP (hence, we must dispense with the theta-criterion, since this is also a theta-position, where the DP is assigned a second theta-role – or, in other version, a complex theta-role);

iii) since Spec,VP is not a Case position, the clitic (the only form available for this subject, since it is co-referent with some DP in the matrix clause) raises to T (which is a position already projected for independent reasons);

iv) now take the unaccusative in (378b): the subject moves from its postverbal position and, as Spec,VP is not projected (the verb has no external theta-role to assign), the clitic raises directly to T;

v) in both – (378b,d) – the subject is stranded in this position, since it is not allowed to move again, to the matrix object (in Müller & Sternefeld 1993 it has been proposed that the movement object-subject-object is not licit);

vi) notice that movement from the embedded Spec,VP position, if the subject has been generated there – which is the case in (378a,c) – is allowed, and that is why we may have a “silent” embedded subject. In fact, what we have is a trace of DP-movement, in the fashion proposed in Hornstein 1999 (in (377) we have the representation of a case of subject “control”, in these Capeverdean sentences we have a case of object “control”).

As for the cases where we have a clitic in the embedded subject position, they are not problematic for this approach, since this clitic surfaces on T, as proposed earlier for all subject clitics; this T is not active, though; the subject clitic is somehow licensed by the preposition pa.
If we assume that these sentences are instances of ditransitive constructions, the *pa-*
clause has the direct object grammatical function. Either way, we know that it selects
for a T complement – and not a VP, as it might seem possible from (378). This can be
shown by the possible embedded negation *ka* and also the affix *-ba*, which are not
prohibited in these contexts, as we have observed and as shown in table (366). Consider
a change of tense for (378c,d):

\[(379)\]
\[
a. N \text{ flaba } Djon p'{-}e / pa (\textit{ka}) \text{ lababa } mo. \quad \text{(transitive)}
\]
\[
1SG \text{ tell.TMA Djon to 3SG / to (NEG) wash.TMA hand}
\]
\[\quad \text{‘I had told Djon (not) to wash his hands.’}\]
\[
b. N \text{ flaba } Djon p'{-}e (ka) \text{ lababa. / * N flaba } Djon \text{ pa (ka) lababa. (reflexive)}\]
\[\quad \text{‘I had told Djon (not) to wash (himself).’}\]

Thus, we may say that *pa* selects for an unactive T which, in the fashion proposed for
full clauses with an active T, accommodates the subject clitic, the negation and the
TMAs that are required by the temporal reference at stake\(^{217}\). The difference in these
particular contexts is that the subject clitic has been moved from the complement
position, has landed on Spec,VP and then has moved again for case reasons, thus also
satisfying the PF requirements regarding adjacency between the TMAs and the verb.

This differs from the proposal in Hornstein 1999 in that this author considers the
existence of an embedded Spec,IP (check the representation in (377)), whereas I
consider that in Capeverdean such position is not projected when not needed.

Finally, let me consider cases of subject “control”, since I have expressed the hope that
these “reflexive” properties could illuminate other instances of embedded non-finites
with “silent” subjects. Verbs such as *expect* or *hope* are not easy to test in Capeverdean,
since my informants resort to some paraphrases that do not include the constructions I
am looking for. But one of the most common is *kre* ‘want’. Let us go back to the
example in (338c), here repeated as (380a); one more relevant example has been added.

\(^{217}\) Recall that the presence of the embedded *-ba* (and also *ta*, in the complements of perceptive verbs)
does not preclude the non-finiteness of T, since it (they) participate in the construction of a temporal
reference that is certainly not independent of the matrix clause.

323
The prohibition of the embedded negation in this context may be a sign that there is no embedded T, but rather this is a complex V head; this hypothesis is borne out by the fact that the embedded -ba is only possible when there is a matrix -ba, which suggests that the matrix T lowers to both V morphemes\textsuperscript{218}.

As for a different “control” verb, we may try \textit{purmete} ‘promise’, in the version where it does not select an embedded complementizer and cannot have an embedded independent tense (it is indeed an embedded infinitive). As we have seen in table (266), it may select for the embedded negation and for the embedded -ba. Notice, however, that this verb is not very common in spontaneous speech, since the verb mostly used in this sense is, as shown earlier, \textit{fla} ‘say’, with the correspondent requirements.

\begin{itemize}
  \item [(381)] \textit{N purmeteba (*N) ka studaba pa testi.}
  \end{itemize}

  ‘I had promised not to study for the exams.’

We may argue, thus, that the verb selects for an embedded T, which is not active but may accommodate \textit{ka} and -\textit{ba}; the embedded subject raises directly to the matrix Spec,VP and then to the matrix T (if it were a full DP, we could say it would raise from the embedded Spec,VP to a matrix Spec,TP that would be created in the course of this derivation; but since T is not able to license the DP, this intermediate movement is unproductive, hence it is not clear why this operation should occur).

In this section I have addressed the question of whether there is a PRO in Capeverdean. Before providing a tentative answer I have discussed two topics, supported by information that has been presented in the previous chapters: i) given what we know regarding the functional structure of the language, namely that the Spec,TP position is

\textsuperscript{218} This context respects the definition of complex predicate as formulated in Baker 1996:338: a complex predicate is “any inflectional domain that contains two distinct morphemes, each of which selects at least one phrasal argument in its theta-grid.” Thus, \textit{kre} selects the DP subject, whereas \textit{ser} selects the DP \textit{nha mos}.\footnote{\textsuperscript{218} This context respects the definition of complex predicate as formulated in Baker 1996:338: a complex predicate is “any inflectional domain that contains two distinct morphemes, each of which selects at least one phrasal argument in its theta-grid.” Thus, \textit{kre} selects the DP subject, whereas \textit{ser} selects the DP \textit{nha mos}.}
only projected when needed, is there any sense in resorting to a much problematic null category only to fill in this position? ii) given what we know about the selectional properties of predicates regarding their infinitival complements, can’t we propose that such a position is not needed, for there is no subject to be spelled out there?

The main conclusion has been that there is no PRO in Capeverdean, at least in the traditional sense of the necessary subject of infinitival clauses. This implied the assumption of Hornstein 1999 proposal of a Movement Theory of Control, which predicts that the null subjects of non-finite clauses are in fact traces of DP movement.

Grounded on evidence provided by some particular Capeverdean contexts – embedded reflexive contexts which have been analyzed in the last chapter as having the subject generated in the complement position –, I have proposed that in embedded infinitives Spec,TP is not projected.

In the cases where the embedded subject has been generated in Spec,VP, it may move to the matrix clause, where it is assigned case (either by the verb, in the complement position – like in ECM –, or by the matrix T). In the cases where the embedded subject has been generated in the complement position (unaccusatives or these particular reflexive entries), it raises for case reasons (just like in the passives).

Whereas the Spec,VP in unaccusatives is not projected (the verb has no external theta-role to assign), in the “reflexive” entries it is projected (recall that these verbs are transitive). Thus, on its way to a case position (supposedly T) the subject lands on Spec,VP, where it is assigned a second theta-role. It raises to the embedded T, from where, just as with the subject of embedded unaccusatives, it is precluded from moving to the matrix clause – according to the proposal in Müller & Sternefeld 1993, the movement object-subject-object is not licit. In these specific embedded contexts where T is not active, the abstract Case requirements of this DP are satisfied by the preposition *pa*, and the derivation converges.

Before addressing these particular questions on Capeverdean, I have discussed the nature of PRO as an idiosyncratic category, briefly reviewing some recent arguments in the literature (6.3.1). In 6.3.2 I have brought into discussion the above mentioned reflexive contexts, in the hope that a possible approach on these would bring some useful predictions to the problem of subjects in null infinitival clauses.
6.4 Conclusion

In this chapter I have discussed embedded predicates that are expected to be infinitival forms. Since there is no specific verbal infinitival morphology in Capeverdean, I have proposed that the diagnostics for non-finiteness be grounded on other levels: i) the syntactic structure of the embedded clause, and mainly ii) the temporal reference contained in these embedded environments.

These questions concerning the visible materials which a non-finite clause may display have been described in 6.1. In section 6.2 the list of questions has been narrowed, focusing mainly on the availability of an independent tense; two distinct studies on infinitival tense (Wurmbrand to appear and Gonçalves 1999) have been summarized, and Capeverdean predicate idiosyncrasies have been discussed. In section 6.3 a different matter on embedded infinitives has been at issue: the subjects that are allowed/obligatory in some of these contexts. A few questions concerning the null category PRO have been listed, and also the steps involved in the alternative approach on control as movement, in a fashion somewhat (although not totally) akin to raising, as proposed in Hornstein 1999 and subsequent work.

The main conclusion has been that, coherently with what has been assumed along this dissertation, Capeverdean Spec,TP is only projected when needed. In some embedded non-finite contexts even T is not projected: the structure is a complex predicate, with a complex V head, allowing for -ba to lower to both V morphemes; this is the case of kre ‘want’. Also manda ‘order’, when it does not select for a pa-clause, allows no embedded material other than -ba (no other TMA, no ka, no subject clitic); manda has, however, an obligatory overt object (Es mandaba mi baba ‘They had ordered me to go’); this fact suggests that there must be a defective embedded T, for the matrix -ba cannot lower to the embedded verb – the necessary adjacency conditions are not met (this is a typical ECM, where the matrix object has been analyzed as the subject of the embedded verb, which has raised for Case reasons); the prohibition of the other materials in the embedded domain is accounted for by semantic determinations.

In the others, where there is indeed an embedded T, we observe that its specifier is not projected, since no DP needs be merged to it. This T is sufficient to accommodate the negation ka, the affix -ba (which then lowers to V), the morpheme ta (with matrix perceptive verbs), and also the subject clitic, whenever it surfaces in the embedded
domain. These latter contexts, where an embedded subject clitic is permitted, are the ones introduced by the complementizer pa, which satisfies the Case needs of the clitic.

As a side note, (382) displays distinct properties of Capeverdean and English regarding null subjects (* means forbidden; \( \checkmark \) means obligatory; ?? means it depends on the selectional properties of the matrix verb).

(382)

<table>
<thead>
<tr>
<th>matrix finite T</th>
<th>embedded non-finite T</th>
</tr>
</thead>
<tbody>
<tr>
<td>referential null subject</td>
<td>referential null subject</td>
</tr>
<tr>
<td>Capeverdean</td>
<td>*</td>
</tr>
<tr>
<td>English</td>
<td>*</td>
</tr>
</tbody>
</table>

The whole comparison with English is interesting, since some parameter values proposed for Capeverdean in this dissertation seem coherent with a common assumption among many creolists (Bickerton 1984, Muysken 1988, a.o.): the unusual circumstances of creole acquisition (/creation) may explain a propensity to unmarked (weak) values. And yet some of these values “typical” of Creoles which are present in Capeverdean are also present in English: the lack of referential null subjects in matrix clauses, lack of V-to-T and SVO order. Naturally, these parametric values are not sufficient to consider a language as a Creole. Neither is this simple enunciation of the parameters of much meaning, for it says nothing about the particular conditions underlying each of them in each language.

Resuming (382), the visible distinction regarding subjects, as seen earlier, concerns expletive finite constructions: English needs an overt expletive, Capeverdean forbids it.

A theoretical question remains open: is there a way of uniformly account for these facts without resorting to the EPP? The values concerning referential null subjects could indeed be accounted for by the abstract property called Case, as long at is related to active T, as is traditionally assumed, but also dependent on the very existence of a DP requiring this feature checking. As for the distinct values regarding expletives, this Case feature seems powerless.

The conclusion is as expected. This must be subject to further research.

\(^{219}\) Except for pa-clauses with predicates whose subject has been generated in P’s complement position.
Concluding remarks

In this dissertation I have described various puzzles concerning Capeverdean predicates. After the Introduction, in Chapter One, where I have addressed some theoretical and methodological issues and summarized the main questions that have guided my research, Chapter Two has been centered on the Capeverdean inflectional domain, and the proposals assumed in my previous works have been developed. Grounded on some distinct theories on temporal reference and supported by extensively described Capeverdean empirical facts, the main proposal has been that the pieces of inflection surface on multiple heads adjoined to T. Besides accommodating all the TMA morphemes, Capeverdean TP is sufficient to account for the negation head *ka*, the copulas *e* (Present) and *era* (Past) and, finally, subject clitics.

In Chapter Three I have introduced some issues concerning the Verb domain, including the parameter values traditionally connected with verbal features. The nature of Case and agreement has been discussed, as well as the role of these in the licensing of DPs. The main claim has been that morphological, rather than abstract features seem to be at stake. Nevertheless, an abstract (Case) feature must be considered, to account for the displacement of subjects from their base position in Capeverdean: subject clitics surface on T, subject DPs surface on Spec,TP; both these pronounced copies respect the adjacency requirements of TMAs regarding the verb. The V-to-T parameter has been addressed as well. Since no overt verbal agreement morphology exists in the language, the main claim has been: there is no empirical reason to posit V-to-T in Capeverdean.

In Chapter Four the description of phenomena inside the Verb domain has been extended, and they have been accounted for on the grounds of the previous proposals, namely that the verb stays in its base position. The approach to these topics required the summary of different joint works developed over the last few years, with some new developments in the spirit of the present work. These specific puzzles have been, respectively: the selection of different object pronominal forms, depending on the presence of postverbal -*ba*; the lack of any anaphoric expression in some reflexive contexts; the structure of ditransitives.
In Chapter Five I have focused on the null subject parameter. After the analysis of relevant data, it has been argued that there is no empirical reason to state that Capeverdean is a pro-drop language: referential *pro* is prohibited in all matrix clauses. In expletive constructions no expletive subject is overtly realized, which, together with the proposals in the previous chapters, has led to the conclusion that, when not needed, Spec,TP is not projected. In order to account for this, the Extended Projection Principle (EPP) has been discussed; the final proposal is that at least in Capeverdean the EPP does not hold. In other words, if active T has a second edge feature in the language this feature says that something can be merged to it (where ‘can’ means ‘may’, not ‘must’): its specifier is only projected whenever some DP already present in the derivation must move to this position (internal merge).

In Chapter Six I have addressed embedded non-finite clauses. First of all, some infinitives have been identified; since there is no specific morphological marking for infinitives I have used different diagnostics: a) the materials allowed in the embedded clause, in order to infer if its T is defective, and b) the impossible independent tense regarding the matrix clause. Subsequently, some hypotheses on their internal structure have been proposed, also accounting for overt subjects displayed in some of them, which are introduced by *pa*. The results of these descriptions have led to the conclusion that these embedded infinitives do not have PRO in their subject position. The embedded subject is assigned one theta-role in its base position and moves afterwards for Case reasons. In their way up, these subjects may be assigned a second theta-role.

These complex facts about the language have inspired the next step in my research program, on language acquisition. Namely, it will be interesting:

i) To determine whether Capeverdean children produce: a) null subjects, in the same stages as children acquiring other “non pro-drop” languages, such as English; b) root infinitives. (Guasti 2002, Hyams 1996, 2005, Rizzi 1995)

ii) To describe the stages of acquisition of TMAs and the construction of temporal reference, and also to investigate any possible relation between the acquisition of the functional domain and any production of null subjects.

I firmly believe that the data gathered in these future acquisition studies will contribute to more accurately describe Capeverdean infinitival contexts and to better understand the proposed values of different parameters, namely on: i) various aspects regarding null subjects; ii) V-to-T and iii) the EPP.
References


Language 63: 741-781.

Languages. In M. DeGraff (ed) Language Creation and Language Change:
Creolization, Diachrony and Development. Cambridge, MA: MIT Press, 329-
373.


M. Butt & T. H. King (eds): Proceedings of the LFG97 Conference,
University of California, San Diego, CSLI Online Publications.

Cabredo Hofherr, P. 1999. Two German impersonal passives and expletive pro. Catalan 

Cardinaletti, A. & M. Starke. 1994. The typology of structural deficiency: on the three 
grammatical classes. Unpublished manuscript, University of Venice and
University of Geneva.

Cardinaletti, A. & M. Starke. 1994. The typology of structural deficiency: a case study 
on the three classes of pronouns. In H. van Riemsdijk (ed) Clitics in the 

adverbs in Portuguese. A. T. Pérez-Leroux & Y. Roberge (eds) Romance 
Linguistics: Theory and Acquisition. Amsterdam/Philadelphia: John 
Benjamins, 95-110.

Castro, A. & F. Pratas. 2006. Capeverdean DP-internal number agreement, additional 
arguments for a Distributed Morphology approach. In J. Costa & C. F. Silva 

Chierchia, G. 1989. A semantics for unaccusatives and its syntactic consequences, ms., 
Cornell University.


and Binding. Cambridge, MA: MIT Press.


Praeger.


Suzuki, M. 1994. The markers in Cape Verdean CP. Unpublished manuscript, CUNY.


Williams, E. 1985. PRO and subject of NP. Natural Language & Linguistic Theory 3:297-315


