Clitics as Morphology

Ana Alexandra Ribeiro Luís

A thesis submitted for the degree of Doctor of Philosophy

Department of Language and Linguistics

University of Essex

March, 2004

Abstract

The purpose of this thesis is to investigate the morphological behaviour of pronominal clitics in European Portuguese (EP) and to develop an inflectional account of cliticisation within the theory of Paradigm-Function Morphology (Stump 2001). It is argued that EP clitic pronouns exhibit most of the affixal properties attested for clitic pronouns in various Romance languages (Auger 1994 and Miller&Sag 1997, for French; Monachesi 1999, for French, Brines 2001, for Spanish). These similarities include rigid ordering, idiosyncratic co-occurrence restrictions, fusion, syncretism, and allomorphic variation. Affixal status is also supported by the behaviour of EP enclitics (i.e., postverbal clitics) which form a morphologically cohering unit with the verb. In particular, enclitics cannot be separated from the verb, may intervene between the verbal stem and tense/agreement suffixes, and induce stem allomorphy. Based on this evidence, clitic sequences are best viewed as affixal sequences and postverbal clitics are best analysed as verbal suffixes.

EP proclitics (i.e., preverbal clitic pronouns) however seem to complicate the inflectional status of the EP clitic system. Even though they are phonologically exactly identical to enclitics, they display distributional and scopal properties that are untypical of verbal affixes: they can have wide scope over two conjoined VPs and do not need to be strictly adjacent to the verb. One crucial question then is how to accommodate the phrasal properties of preverbal clitics within an inflectional account of cliticisation. Previous studies have argued that the scopal and distributional properties completely invalidate an inflectional analysis, claiming that EP clitic pronouns must therefore be regarded either as special word-level units (Vigário 1999b, Gerlach 2001a) or as morphosyntactic 'hybrids' (Crysmann 2002).

This thesis draws on the concept of phrasal affixation, developed by Klavans (1985), Anderson (1992) and Spencer (2000), and argues that proclitics constitute genuine phrasal affixes. Given this assumption, the asymmetry between proclitics and enclitics is captured as a difference in status between the word-level and phrase-level placement of the same affixal unit: while proclitic prefixes attach to a phrasal node, enclitic suffixes combine with a morphological base. Formally, an analysis is developed which assigns affixes the ability to be positioned either as verbal suffixes or as phrasal prefixes.

Contents

Acknowledgments

This work has received support and motivation from a number of people over the years. I would like to begin by thanking Andrew Spencer, my thesis supervisor, for his friendship and continuous encouragement. Andrew supervised my MA dissertation on the morphology of English auxiliaries in 1997, and convinced me to look at clitics in my own language for my PhD. Thanks to his genuine enthusiasm for linguistics and morphology, I had the privilege of sharing with him lengthy hours of fruitful discussion. They helped shape my views on cliticisation and sharpened my approach to language. The many meetings we spent discussing the theory of Paradigm-Function Morphology, which at that time was only in circulation as a manuscript, were fundamental for my research, and his contributions during the various stages of my research were a constant source of inspiration. I also thank Louisa Sadler, my co-supervisor, for introducing me to Lexical-Functional Grammar and discussing with me those aspects of the data more directly related to syntax. Louisa guided me through the complexities of the morphology-syntax interface and discussed with me many interesting ideas, some of which could not be included in this thesis for reasons of time and space. They did however help me widen my horizons about cliticisation and grammar.

My work has benefited equally from the teaching I have gained at the University of Essex, with my teachers Keith Brown, Doug Arnold, Andrew Radford, Iggy Roca and Martin Atkinson. At the University of Coimbra, I must thank, for my earliest studies in Linguistics while still an undergraduate, my teachers Maria Francisca Athayde, Adelaide Chichorro, the late Maria Manuela Rocheta Santos, Penelope Stonehouse David, Mark Hayman, Isabel Pereira and Maria José Moura Santos.

I would also like to express my gratitude for the many diverse contributions from colleagues and linguists. In the early stages of my research, I benefited from the suggestions put forward by Joan Bresnan, Peter Sells, Jonas Kuhn, Annette Frank, Miriam Butt, Farrell Ackerman, Gert

Webelhuth, Greg Stump, Wyn Johnson, and Graça Rio-Torto, all of whom helped me to resolve some of the initial difficulties I encountered. My appreciation also goes to the audiences of the conferences in which various parts of this thesis were presented, in particular Eulalia Bonet, David Bobaljik, John McCarthy, Danielle Goddard, Doug Arnold, Janet de Cesaris, Dan Brassil, John Frompton. I am also indebted to a great number of people for sending me bibliographic material and, in some cases, for furnishing me with copies of work still in progress. It is for such generous assistance that I thank Julie Auger, David Heap, Miriam Mithun, Paola Beninca, Enric Vallduvi, Paul Kiparski, Phoevos Panagiotidis, Natalia Brines, Birgit Gerlach, Marina Vigário, Berthold Crysmann, Alice Harris, Ana Maria Martins, Isabel Pereira, Mark Janse.

It was with the help of my Head of Department of the Department of Anglo-American Studies at the University of Coimbra, Martin Kayman, and my Pedagogical Supervisor, Abílio Hernandez Cardoso, that I was able to apply for my PhD course at the University of Essex. With their support, I could enjoy three consecutive years of study-leave without which I could not have benefited from the excellent research environment provided by the University of Essex. Many thanks also to the subsequent Heads of Department, in particular, to Abílio Hernandez Cardoso, Maria Irene Ramalho and Jacinta Matos, who over the years supported my study-leave. I also wish to express my gratitude to my colleagues, Clara Keating, John Holm and Isabel Pereira for taking on extra work to cover my teaching hours during my absence from the University of Coimbra.

Special thanks go to the various entities from which I received financial help. In particular, the A.H.R.B./British Academy and the Portuguese Foundation for Science and Tecnology/Ministry of Science and Higher Eduation for the four-year scholarships which enabled me to stay in residence at the University of Essex during most of my PhD period. I also thank the Luso-American Foundation for Development and the Calouste Gulbenkian Foundation in Portugal, whose grants allowed me to complete a period of study and investigation at Stanford University, with the help and support of Professor Andrew Mannings. Furthermore, I am most grateful to the Faculty of Letters of the

University of Coimbra for financial assistance to attend international conferences and to make essential visits to a number of research centres.

I would also like to thank my friends at Essex, Cristobal Lonzano, Maria Flouraki, Misako Ono, Russell Norton, Valia Kordoni, Natalia Brines, Joseph Gwnebith and Andy Way. Guergana Popova came to Essex in my third year and Ryo Otoguro in my fourth year. I wish we had met much earlier. I thank both for the many discussions about morphology, clitics and morphosyntax. Closer to home, I am grateful to my friends Ana Paula, Osvaldo, Rosa and, in particular, Abílio, for giving me, from time to time, the extra boost I needed to keep going.

Finally, my warmest thanks must go to my parents, parents-in-law and, most of all, to my husband, Fernando Oliveira.

Abbreviations

1	first person
2	second person
3	third person
ACC	accusative
CLASS	conjugation class
CL	clitic
COND	conditional
C-STRUCTURE	constituent structure
DAT	dative
FEM	feminine
F-STRUCTURE	functional strtucture
FUT	future
INF	infinitive
INFL	inflected
IMP	imperative
IMPERF	imperfect
IND	indicative
MASC	masculine
M-FEATURES	morphological features
M-P	morphophonological rules
NUM	number
OBJ	object
PERF	perfect
PER	person
PL	plural
PF	Paradigm Function
PRS	present
PWD	phonological word
REFL	reflexive
RR	realisation rules
SG	singular
SUBJ	subject
TNS	tense

Glosses

Verbal endings (such as tense, person and number suffixes) are separated by hyphens, both in the original language (first line) and in the English glosses (second line):

(i) lev -á -va -mos
take -class1 -Impf. -1.pl
‘(we) took’

Following the Portuguese writing convention, clitic pronouns are separated by a hyphen in potsverbal position (cf. ii) but not preverbally (cf. iii).

(ii) bebi -o
drank -acc.2sg.masc
'I drank it'

(iii) não o bebi
not acc.2sg.masc drank
'I didn't drink it'

Whenever morpheme-by-morpheme glossing is not necessary, word-by-word glosses are adopted. In this case, each word in the original language is lined up with the corresponding English meaning:

(iv) Quem está na rua?
who is in.the street
'Who is in the street?'

Part I - Introduction

PART I presents the empirical and theoretical assumptions underlying the thesis.

CHAPTER 1 provides an outline of the main research questions and surveys the linguistic phenomena under discussion. This is followed by a short description of the Portuguese clitic inventory and an overview of the structure of the thesis.

CHAPTER 2 introduces the theory of Paradigm-Function Morphology as developed by Stump (1993, 2001) and discusses the main theoretical assumptions. Emphasis is given to the interaction between realisation rules and paradigm functions: while realisation rules derive affixes as the phonological output of morphosyntactic feature sets, paradigm functions determine the complete set of inflected forms for a given lexeme. More recent developments of the theory are addressed in Parts II and III of the thesis.

CHAPTER 3 introduces the phenomenon of Romance cliticisation and surveys some of the claims made in the literature about the grammatical status of pronominal clitics. I shall compare two different 'morphological' approaches to cliticisation and survey previous studies on EP clitic pronouns.

Chapter 1 Main research questions

Recent studies have shown that European Portuguese (EP) clitic pronouns, as given in (1), display numerous non-syntactic properties.

(1) a. A Joana apresentava *-se* *lhes*.
the J. introduced- refl.3sg dat.3.pl
'J. introduced herself to them'

b. A Joana mostrar- *lhes* -á os documentos.
the J. show- dat.3.pl- Fut.3sg the documents
'J. will show them the documents'

Authors such as Vigário (1999b), Gerlach (2001a) and Crysmann (2002) have argued that clitic pronouns cannot be adequately accounted for under a purely syntactic approach. Vigário (1999b), for example, examines non-productive shape alternations within the theory of Precompiled Phonology (Hayes 1990) and derives clitics as phrasal allomorphs. Gerlach (2001a), on the other hand, looks predominantly at the idiosyncrasies of clitic combinations and argues that clitic clusters should be derived through morpho-lexical constraints, within an OT model of grammar. Crysmann (2002) analyses the distribution of enclitics and proclitics, within linearisation-based Head-Driven Phrase-Structure Grammar (Kathol 2000), and takes the view that EP clitic pronouns constitute morphosyntactic hybrids, generated by the morphology and positioned by the syntax.

This thesis also takes the position that cliticisation in EP cannot be regarded as a purely syntactic phenomenon. However, a different proposal shall be made to account for the numerous non-syntactic properties displayed by the data. My goal will be to argue that clitic pronouns constitute, in

effect, inflectional affixes, and to the EP clitic system as an inherently inflectional system[1]. Based on criteria previously used for determining the inflectional status of clitic pronouns in Romance, I shall motivate this claim with empirical evidence, based on rigid ordering, idiosyncratic co-occurrence restrictions, fusion, syncretism, allomorphic variation, and stem allomorphy, among other properties. This approach will be in line with the work of numerous linguists who, based on similar criteria, have recently argued in favour of an inflectional approach to clitic pronouns in Italian, French and Spanish (e.g., Monachesi 1999, Miller 1992, Auger 1994 and Miller&Sag 1997, Brines 2001, Nishida 1987).

Elaborating on previous claims by Zwicky (1987), Stolz (1992) and (Spencer 1991) about the affixal natures of EP enclitics, the behaviour of clitic pronouns shall be captured within the realisational theory of Paradigm-Function Morphology (Stump 2001). One crucial aim of this thesis will be to argue that preverbal clitics (i.e., proclitics) also constitute inflectional affixes, albeit of a different type. Exploring insights by Klavans (1980), Anderson (1992), Spencer (2000) and Legendre (2000b), it is argued that EP occupies a unique place in Romance in that proclitics constitute, in effect, phrasal affixes.

In the following sections, an overview of the main research questions is offered (1.1), the data set investigated in this study is introduced (1.2) and the organisation of the thesis is laid out (1.3).

1. Overview

The following subsections introduce the topics addressed in this study and survey the main issues they raise. The topics include well-known phenomena such as the preverbal/ postverbal position of pronominal clitics and the internal structure of clitic clusters; they also include more language-

[1] In this study, therefore, the term 'clitic' is used as a purely descriptive cover term. Formally, 'clitic

specific phenomena such as the word-internal placement of clitic suffixes in EP, the asymmetry between preverbal and postverbal placement of clitics, and the phrasal properties of pronominal clitics.

1.1.1 Cliticisation as inflectional morphology

The first step in this study will be to argue that clitic pronouns, both in verb-final and verb-internal position, as illustrated in (1a) and (1b), are best analysed as verbal suffixes (Zwicky 1987, Stolz 1992, Spencer 1992, Luís 2003a,b).

1.1.1.1 Clitic suffixes

Enclitics appear verb-finally after tense and agreement suffixes, as in (1a), and constitute the default case in EP. Phonological, morphological, lexical and syntactic criteria support the claim that enclitics behave like suffixes: they cannot be separated from the verb, they cannot have wide scope over coordination, they undergo allomorphic alternations and induce stem allomorphy. In addition, multiple sequences of clitics exhibit rigid ordering, co-occurrence-restrictions, clitic allomorphy, portmanteau forms and syncretism. Given this evidence, it is claimed that affixal properties pose serious problems to a syntactic account but fall out naturally if enclitics are analysed as part of verbal morphology (Stolz 1992, Zwicky 1987, Spencer 1992, Crysmann 2002, Luís 2003a, to appear). This position reflects similar claims that have been made for pronominal clitics in other Romance languages (e.g., Monachesi 1999; Auger 1994, Miller 1992, Miller&Sag1997, Brines 2001).

Formally, cliticisation raises interesting conceptual questions about the format of realisation rules, the organization of clitic rules with respect to ordinary inflectional rules, and the structure of

---

pronouns' will be regarded as affixes.

verbal paradigms. To account for the cohering relation between the verb and the enclitic, I suggest that the verb-enclitic unit constitutes an 'extended' verb form comprising both a sequence of ordinary suffixes and a sequence of clitic suffixes, with pronominal clitics forming the outer layer of verbal inflection (see also Anderson 1992, Halpern 1995, Spencer 2000, Legendre 1996, 1998, 2000a). How to explicitly account for cliticisation within inflectional morphology is at the center of this thesis (cf. Chapters 6, 7 and 9), and shall be explored within the realisational theory of Paradigm-Function Morphology (Stump 1993, 2001; Spencer 2000).

1.1.1.2 Verb-internal suffixation

Clitic suffixes have the ability to appear word-internally, i.e. between the verb stem and the tense/agreement marker, as illustrated in (1b). Even though the phenomenon is well-known in the literature (Spencer 1992, Harris&Vincent 1988, Halpern 1995, Zwicky 1987), no explicit analysis of the data has yet been proposed. One of the aims of this thesis shall be to accommodate mesoclisis as an inflectional phenomenon, following insights from Spencer (1992) and Zwicky (1987). It will be argued, contrary to syntactic analyses, that Portuguese verb forms with verb-internal clitics constitute inflected words and not analytic verb forms (Zwicky 1987, Spencer 1992, Stolz 1992, Crysmann 2002). To support this view, I shall draw on genuine affixal properties of verb-internal clitic suffixes, on the one hand, and the affixal behaviour of the tense and agreement markers, on the other.

The purely inflectional nature of mesoclisis is captured by analyzing the verb-internal placement of clitic suffixes as an instance of affix metathesis. It is assumed that the sequence of rule blocks realising clitic affixes applies before the sequence of rule blocks deriving ordinary tense and agreement suffixes. From this viewpoint, enclisis is derived as a case of suffixation to an inflected verb form (cf. 1a) and mesoclisis is derived as suffixation to an infinitival stem (cf. 1b). Conceptually, 'mesoclisis' sheds new light on standard views about the inflectional attachment of

clitics to a verbal base by challenging the widely-held assumption that affixal clitics in Romance attach to fully inflected verb forms (as assumed in the inflectional studies of Anderson 1996, Monachesi 1999, Miller&Sag 1997); instead, the data shows that clitics can also appear as 'inner' layers of inflection, as in (1b), where clitic suffixes occur before the tense and agreement markers. The ability for 'inner' and 'outer' layers of affixation to appear in reversed order will be captured within an extended version of Paradigm-Function Morphology, proposed initially by Spencer (2000) and further developed in Luís (2003b, to appear), Luís&Spencer (in press).

1.1.1.3 Preverbal clitics

One further problem raised by EP clitic pronouns concerns the unequal status of enclisis and proclisis (Duarte et al. 1995, Vigário 1999b, Crysmann 1997). While enclitics behave in all respects like verbal suffixes (1.1.1.1), proclitics exhibit phrasal behaviour: they may be separated from the verb by single words (Martins 1994), and may be shared over coordinated verb phrases (Matos 1997), as shown in (2). Based on this evidence, the crucial question is how to accommodate the distributional and scopal properties of proclitics without weakening the claim that clitics in EP constitute inflectional affixes.

(2) a. Acho que o João lhe ainda não revelou o segredo.
think.1sg that the J. dat.3.sg yet not revealed the secret
'I think that J. has not yet revealed the secret to him/her'

b. Todos os amigos me escreveram postais e enviaram prendas.
all the friends dat.2sg wrote cards and sent presents
'All my friends wrote me cards and sent me presents'

The syntactic behaviour of proclitics has been used by Vigário (1999b), Gerlach (2001a) and Crysman (1997, 2002) to rule out the purely inflectional status of clitic pronouns. Contrary to these studies, I argue that upon closer inspection enclitics and proclitics should be regarded as the same

affixal unit. Supporting this view is, most of all, the fact that proclitics are formally exactly identical to enclitics and that proclitic clusters exhibit exactly the same range of allomorphic and ordering effects. The homophony and exact morphotactic structure suggests that enclitics and proclitics constitute the same affixal units, being generated by the same inflectional mechanisms. To account for the distributional and scopal differences between enclitics and proclitics, shown in (2), it is argued that proclitics constitute *phrasal* affixes: they attach to a phrasal node and do not form a morphologically cohering unit the adjacent host, unlike enclitics. Under this view, the asymmetry is captured at the level of clitic placement. It is therefore argued that the phrasal behaviour of proclitics does not invalidate the affixal status of pronominal clitics in EP, contrary to previous claims by Vigário (1999b), Crysmann (1997) and Gerlach (2001).

The notion of phrasal affix has been introduced by Klavans (1980, 1985) to explain why certain affixal clitics attach to phrasal positions rather than to hosts belonging to a specific category. Typologically, we find that affixal clitics are generally either uniformly attached in the morphology (as morphological suffixes and prefixes, as assumed for Italian by Monachesi 1999, and French by Miller&Sag 1997); or uniformly of a phrasal kind, as assumed for Serbo-Croatian by Anderson (1996) and for Bulgarian, by Spencer (2000) and Legendre (1996, 2000b). For EP, the claim is that proclisis is an instance of phrasal affixation while enclisis is best analysed as morphological suffixation. Even though asymmetric placement is somewhat marked within clitic systems, it is by no way unique to EP (cf. Chapter 8 for cross-linguistic evidence).

1.1.1.4 Dual placement

Regardless of whether affixal clitics attach within a phrasal or morphological domain, an inflectional approach to cliticisation must be able to account for the 'dual' placement of clitics, i.e., their ability to appear before or after a given host. This question is not only relevant for clitic

systems (in most Romance and Slavic languages) but also for ordinary inflectional systems, as illustrated by Stump (1993), for Fula and Swahili, and by Noyer (1994), for Huave and Afar.

Crucial in accounting for 'dual' placement is the idea that the formative (i.e., affix) appearing in preverbal position is phonologically exactly identical to the affix appearing postverbally. In this respect, we are not in the presence of two affixes but of just one and the same affix. Dealing with these issues for Fula, Stump (1993) suggested that realisation rules deriving 'dual' affixes should not specificy the direction of attachment of the exponent. This proposal effectively entails a separation between *realisation rules* and *placement rules*: while the former derive affixal clitics as pure form, the latter determine whether the affix appear preverbally or postverbally. Stump's 'separation' hypothesis provides an insighful basis for deriving enclitic and proclitic placement of a given clitic inventory. However to account for clitic placement in EP, placement rules will need to specifiy two types of information: not only the direction of attachment (as in Stump's proposal), but also the identity of the host and the domain of attachment. As to the direction of attachment, clitics are either prefixal or suffixal; as to the domain, they attach either to a verbal stem (as EP enclitics; or Spanish enclitic and proclitics) and to a verbal node (for EP proclitics).

It is worth emphasizing that this model of clitic placement, not only derives cases of asymmetric placement, as in EP, but also the more common cases of uniform placement, either within the morphology (e.g. Miller&Sag 1997, for French) or within the syntax (Anderson 1992, for Serbo-Croatian, Kawakwala).

1.1.2 Clitic clusters

One further goal of this thesis will be to focus on clitic sequences and to examine how they behave both internally (with respect to each other) and externally (with respect to the host).

1.1.2.2 Internal idiosyncrasies

Starting with the behaviour of clitics inside the clusters, EP is different form Romance languages such as Italian and French in that only two types of clitic sequences are attested: the reflexive-dative (3a) and the dative-accusative (3b) patterns. Otherwise, the affixal regularities observed in EP are very identical to those found in other Romance languages:

(3) a. O João apresentou -se -lhe.
the J. introduced -ref.3sg -dat.3.sg
'J. introduced himself to him/her'

b. A Maria e a Joana pediram -vo -lo ontem.
The M. and the J. asked.for -dat.2pl -acc.3sg.masc yesterday
'Maria and Joana asked you for it yesterday'

fixed order and co-occurrence restrictions are two recurrent properties of clusters in EP and other Romance languages which contrast sharply with the combinatorial possibilities available for nominal phrases in EP (Perlmutter 1970, Simpson&Withgott 1986). These irregularities do not appear to be motivated by any obvious semantic or syntactic principles given that most of the ungrammatical combinations are acceptable with free pronominals and full noun phrases. In addition, other cluster-internal aspects such as a) allomorphic alternations, b) portmanteau forms and c) syncretism also seriously challenge the view that clitic pronouns are syntactic units. It shall therefore be argued, in line with previous studies on Romance and Slavic, that morphophonological effects and restrictions on the combination of clitics must be derived through inflectional mechanisms (Miller&Sag 1997, Anderson 1996, Legendre 2000a, Spencer 2000). Formally, it is argued that ordered rule blocks within Paradigm-Function Morphology (Stump 2001) nicely generate these idiosyncratic clitic sequences (Luís 2001a, 2003b).

1.1.2.3 Invariant order

We finally examine the linear order of clitics and examine why clitic order is not affected by preverbal and postverbal placement (4).

(4) a. O Manuel apresentou -se -lhe.
the M. introduced -refl.3sg -dat.3sg
'Manuel introduced himself to him/her'

b. Até o Manuel se lhe apresentou.
even the M. -refl.3sg -dat.3sg introduced
'Even Manuel introduced himself to him/her'

By default, affixes nearer the root in postverbal position should also be nearer the root in preverbal position. However, as (4) illustrates, the clitic clusters seems to move as a whole unit. Similar observations have been made by non-syntactic approaches to clitic clusters (Gerlach 1998, Anderson 1995, Spencer 2000, Luís 2001a). In particular, Gerlach (2001) has shown that clitic clusters should be derived as whole units prior to their placement in proclitic or enclitic position. In this thesis, invariant clitic order shall be captured through the separation between rules of exponence and rules of placement (cf. 1.1.1.4). In particular, it will be assumed that realisation rules derive clitic clusters as an inflectional unit, while placement rules determine the preverbal or postverbal position of the cluster, thus captureing the fact that the same clitic cluster appears in enclitic and proclitic position. The analysis, although based on Stump's 'separation' hypothesis, introduces an extension to the theory by enabling placement rules to apply to sequences of affixes rather than to each individual affix.

1.1.3 Final remarks

The above summary shows that the main goal of this dissertation is to explore the affixal properties of clitic pronouns and to develop an inflectional analysis of the data within the realisational theory

of Paradigm-Function Morphology (Stump 2001). The analysis provides a unified account of three crucial aspects about EP cliticisation: the invariant ordering of clitics, the dual placement of clitics and asymmetric placement.

In addition to a purely inflectional treatment of the data, this study will also address some topics on the syntax of cliticisation, in particular, the syntactic conditioning of proclisis. Attention will be paid to the fact that proclisis is not dependent on the finiteness of the verb, as in French, Italian and Spanish, but determined by a set of preverbal triggers. Even though this issue constitutes effectively a topic for further research, it will be addressed mainly to argue that the syntactic factors triggering proclisis do not weaken the inflectional status of the EP clitic system (contrary to claims by Vigário 1999b, Gerlach 2001a, and Crysmann 2002). Formally, I will explore the morphology-syntax interface by placing Paradigm-Function Morphology in correspondence with Lexical-Functional Syntax (Kaplan&Bresnan 1982, Bresnan 2001). The incorporation of realisational morphology withing the LFG architecture follows insights from Sadler&Spencer (2001), Sadler&Nordlinger (to appear), Sadler&Nordlinger (2003), Luís&Sadler (2003), and Otoguro (2003).

1.2 The data set

Throughout this dissertation, European Portuguese (EP) refers to contemporary standard Portuguese spoken in Portugal. Examples are taken from published work and spoken conversation, or provided by myself. In the latter case, my own examples have been frequently checked against the intuitions of other native speakers[2]/[3].

[2] Other varieties of Portuguese are not addressed; for a descriptive overview of pronominal clitics in Brazilian Portuguese, see for example, Cunha&Cintra (1987), Harris&Vincent (1988) and Bechara (2000), and references therein.

[3] My Portuguese variety is spoken in (the northern part of) central Portugal.

Empirically, the goal was to cover a wide a range of clitic patterns - from clitic clusters and clitic-verb combinations to clitic allomorphs – and to examine the implications of this diversity for inflectional morphology. Therefore, notions such as regional variation, style or register do not play a relevant role in the discussion. However, for completeness, it may be worth noting that some of the clitic structures addressed in this study seem to be more typical of certain varieties of EP and some other appear to be more productive in formal contexts. For example, subject to regional variation is the phenomenon of interpolation, which I refer to in Chapter 8. It has been previously regarded as a property of EP (Martins 1994) but it seems to be more productive in the northern variety of Portuguese (Barbosa 1997). In addition, the phenomenon of 'mesoclisis' (i.e., verb-internal placement of postverbal clitics) is more productively used by educated speakers of Portuguese (Duarte et al 1995). Also, the full range of clitic sequences which I consider are subject to style and register: dative-accusative clusters *no-lo* '1.dat.pl-3.acc.sg.masc' and *vo-lo* '2.dat.pl-3.acc.sg.masc' (regardless of gender and number) are mostly used in formal written discourse. It is not clear why this specific combination is avoided, since other two-position clusters appear productively both in formal and informal EP, such as *se-me* '3.refl-1.dat.sg' and *se-lhe* '3.refl-3dat.sg' (cf. Chapter 4). Even though I do not focus directly on the importance of these stylistic or regional differences, the inflectional analysis developed in this thesis can straightforwardly explain why mesoclisis is used less productively and it can also account for the restricted use of interpolation in the less conservative varieties of Portugal.

Finally, one very important property about the use of clitics in contemporary EP is the growing tendency for speakers of EP to use enclitics in proclitic contexts. Close attention shows that the misplacement of enclitics is quite recurrent in the spoken and written discourse of mostly young educated speakers and frequent in colloquial styles of writing, such as sports newspapers, interviews and tv-subtitles. In this thesis, the fact that enclitics are misplaced in informal and spontaneous discourse will not be subject of analysis, but it plays a crucial role in determining the inherently

inflectional nature of the EP clitic system. Because is supports the claim, previously made by Duarte et al. (1995), that enclisis constitutes the default placement in EP (cf. Chapter 3)[4].

I shall now briefly introduce the inventory of clitic pronouns. The complete paradigm is given below:

(5) Paradigm of affixal clitic in European Portuguese (EP)

	Refl	Dat	Acc
1sg	me [mə]	me [mə]	me [mə]
2sg	te [tə]	te [tə]	te [tə]
3sg	se [sə]	lhe [ʎə]	o [u], a [ɐ]
1pl	nos [nuš]	nos [nuš]	nos [nuš]
2pl	vos [vuš]	vos [vuš]	vos [vuš]
3pl	se [sə]	lhes [ʎəš]	os [uš], as [ɐš]

The table shows that EP comprises a rather small inventory of forms. The forms are restricted to accusative, dative and reflexive clitics. This constitutes the 'minimal' inventory of clitics (Gerlach 2001a) in contrast to the larger number of clitics found in Italian and French where 'object' clitics are combined with partitive and locative clitics, and in Romanian where there are also auxiliary clitics. As to the formal make-up of the forms, we notice that whereas 1st and 2nd person singular and plural forms have an identical shape throughout the three cases, 3rd person clitics are completely distinct. Indeed, 3rd person clitics exhibit the highest degree of differentiation: a) number contrast is found in accusative and dative clitics (e.g., compare *o*/*a* with *os*/*as* and *lhe* with *lhes*) and b) gender distinction is available in the plural and singular accusative forms (e.g., compare *o* with *a* and *os* with *as*). On the contrary, 3rd person reflexive clitics are neither marked for gender nor number.

[4] Duarte et al. (1995) supports the claim with data on language acquisition and adult spoken language. As part of my research, I have collected an informal corpus with instances of 'misplaced' enclitics found both in the spoken but also in the written language. For reasons of space, I have not included the data in this thesis.

It is crucial to bear in mind that even though the pronominal forms given in (5) are presented in tabular form, they do not constitute word units. Neither should the table in (5) be viewed as a list of lexical entries for affixes. Even though they constitute affixes, the theory of inflectional morphology adopted in this thesis, namely Paradigm-Function Morphology (Stump 2001), does not treat affixes as listed items in the lexicon but as the phonological output of the feature set realised by that affix (cf. Chapter 2). Thus, the table is only used for expository convenience only.

1.3 Organisation

This section lays out the structure of the thesis. CHAPTER 2 introduces the inflectional theory of morphology which provides the framework within which the data will be analysed. CHAPTER 3 discusses some of the prevailing views about the morphological properties of clitic systems. The empirical data motivating the inflectional approach to clitic pronouns is provided in CHAPTER 4 based on a set of well-known criteria. CHAPTER 5 then provides phonological evidence to further motivate the need for an inflectional analysis of cliticisation in EP. Having provided the necessary morphological, phonological and syntactic evidence, CHAPTER 6 sketches a preliminary approach to enclisis and mesoclisis within Paradigm-Function Morphology, and CHAPTER 7 extends the inflectional analysis to clitic-induced morphophonological effects. In CHAPTER 8, I finally turn to proclisis and examine their phrasal properties. The main goal will be to argue that proclitics should be regarded as phrasal affixes. After examining preverbal clitics, CHAPTER 9 finally provides a comprehensive and unified inflectional analysis of the EP clitic system. Among other issues, I examine properties such as 'dual' placement and invariant clitic order. CHAPTER 10 addresses the morphosyntax of cliticisation, focusing on the syntactic conditioning of clitic placement, within the architecture of Lexical-Functional Grammar. To conclude, CHAPTER 11 summarises the main findings of the thesis. APPENDIX A provides a fragment of the EP verbal paradigm and APPENDIX B

offers a summary of the inflectional rules for EP clitic pronouns. Finally, partial overviews of the thesis are provided at the beginning of Parts I , II and III.

Chapter 2 Theoretical background

This chapter presents the core ideas and key features of the realisational theory of Paradigm-Function Morphology (PFM) as proposed in Stump (1991, 1992, 1993, 1997, 2001). It also provides a broad outline of the architecture of grammar adopted in the thesis, clarifying how the morphology interacts with other modules of grammar.

2.1 Paradigm-Function Morphology

The theory of Paradigm-Function Morphology offers one of the most carefully worked out models of inferential-realisational morphology. The key idea of inferential-realisational theories is that a lexeme´s association with a particular set of morphological features determines the inflectional rules deriving that word's inflectional form. This means that we need to know the full feature set associated to a given lexeme before deciding which affixes realise those features. Affixes are thus regarded as the realisation of the features associated to that word. Other realisation theories include Matthews' (1972) Word-and-paradigm model and Anderson's (1992) A-Morphous Morphology.

Unlike realisational models, morpheme-based (or morphemic) models of morphology are based on the assumption that a) affixes are (sub-)lexical entries which contribute their own meaning and that b) words are formed through the concatenation of affixes. Substantial evidence however shows that there is insufficient evidence for a morphemic view of word-structure. As to a), there is a range of morphological possibilities in natural languages that cannot be properly represented as the addition of an affix; some of these include apophony (i.e., vowel mutation), subtraction (i.e. deletion of phonological material) or metathesis (i.e., the re-arrangement of an affix). As to b), the relation between meaning and form is often many-to-many rather than one-to-one: 'empty exponents' are

formatives which fail to be associated with any morphemic content; 'cumulative exponents' are associated simultaneously with more than one meaning within a given word; 'extended exponence' occurs when one meaning is spread over more than one affix and 'zero exponence' when one meaning has no formal representation. In all these cases, the one-to-one relation between meaning and form, which is essential to the classical morpheme view, fails to occur. Within realisation models, both the more simple cases of agglutinative morphology and the deviations from strict agglutination can be naturally accounted for (cf. Spencer 2000, 2001, ms, for detailed illustration of the shortcomings of the morphemic view of morphology).

The next section introduces the relevant parts of Paradigm-Function Morphology, as proposed in Stump (2001). Three types of rules are primary for the theory: paradigm functions, realisation rules, and morphological metageneralisations. Of these, I shall first refer to realisation rules (RRs) and paradigm functions (PFs), and postpone the discussion about metageneralisations until Chapter 7.

2.1.1 The notion of paradigm and paradigm function

2.1.1.1 Paradigm cells

Paradigm-Function Morphology (PFM) is based on the fundamental assumption that paradigms are central to the definition of the inflectional system of language and are not just epiphenomenal as often assumed in other theories. This centrality is manifested in the assumption that every lexeme has a paradigm and that each inflected form of that lexeme constitutes a cell in that lexeme's paradigm.

The inflectional paradigm of a lexeme L consists of the set of pairs <Y, σ> such that σ is a complete set of morphosyntactic features for L, and Y is a well-formed word realizing both the lexeme L and the set σ. The set of pairs constitute what is commonly known as the cells of the paradigm. In German, for example, the form *Buches* occupies the genitive singular cell, represented as <*Buches*, σ> in the inflectional paradigm of the lexeme BUCH; likewise, the Bulgarian verb

form *kradém* 'we steal' occupies the first-person plural present indicative cell < *kradém*, σ> in KRAD's paradigm. Crucial to the notion of paradigm is the idea that the paradigm of a lexeme L is determined by a) the morphosyntactic features for which L may be specified and b) any restrictions regulating the co-occurrence of such feature specifications. Morphosyntactic features can be of type 'voice', 'tense', 'mood', 'number', 'person'; and the values of these features can be 'active'/'passive' for 'voice'; 'present'/ 'imperfect'/ 'aorist' for 'tense'; 'indicative'/'imperative' for 'mood', etc. (Stump 2001, Chapter 2). Given that a Bulgarian verb cannot be associated with just any combination of these morphosyntactic properties, a set of co-occurrence restrictions is assumed specifying the well-formed sets (cf. Stump 2001, section 2.3 for co-occurrence restrictions and permissible feature combinations in Bulgarian).

2.1.1.2 Defining word forms

The formal device by which each inflected form of a lexeme is derived is called a paradigm function (PF). A paradigm function is a function which takes as input the root of a lexeme paired with a set of morphological properties and delivers the word form occupying the corresponding cell in that lexeme's paradigm. For example, the Bulgarian paradigm function applies to the pairing <*krad*-, σ>, where *krad*- is the root and σ is the complete feature set {Vform:finite, Voice:active, Tense:pres, Mood:ind, Per:1, Num:pl}> to yield the corresponding cell in the paradigm of KRAD, namely <kradém, σ>. The format adopted for the representation of a paradigm function is given in (1).

(1) PF (<X, σ>) $=_{\text{def}}$ <Y, σ>

(1) says that the PF applies to the pair (<X,σ>), where X is the root of a lexeme and and where Y constitutes the inflected form. Also, σ denotes the complete set of morphological features realised by Y. An example of a PF is given in (2), where the root *krad*- is mapped onto a permissible feature set yielding the verb form *kradém*.

(2) PF (<*kra*-,{VFORM:finite, VOICE:active, TENSE:pres, MOOD: ind, PER:1, NUM:pl}>)

$=_{def}$ <*kradém*,σ>

Thus the crucial role of paradigm functions is to derive each inflected form of a lexeme as a cell in that lexeme's paradigm. An important observation for the present discussion is also the fact that each cell constitutes a morphologically complex form whose internal parts cannot be accessed by rules of syntax. Instead, morphological cells must be represented as one syntactic node. The morphological integrity displayed by inflected words follows from the fact that word forms are derived in the morphology (see 2.2 for summary of assumptions about the grammar).

2.1.2 Realisation rules

2.1.2.1 Rules of exponence

Whereas a paradigm function is a function which maps a root of a lexeme onto one of the fully inflected words in the paradigm of that lexeme, realisation rules (RRs) are individual operations which express subsets of morphosyntactic features[5].

As alluded to above, one of the crucial ideas of a realisational theory of morphology is that the morphosyntactic features associated to a given lexeme trigger the application of a set of operations which give you the inflected form of that lexeme. In PFM, these operations are called realisation rules. They apply to roots or stems and realise the morphosyntactic features through distinct markings. In this sense, each marking or exponent (i.e., affix) is the expression of a given set of

[5] To be more precise, realisation rules are of two types: (i) rules of exponence directly specify the concrete exponents associated with the property set being realised; (ii) rules of referral which refer the realisation of some property set to some other realisation rule(s). I will here refer only to the realisations rules in the sense of inflectional rules of exponence (cf. Stump 2001, Chapters 2 and 7, on rules of referral).

features, as illustrated in (3) with the Portuguese verb form *bebíamos* 'we drank'. Each inflectional rule (also known as realisation rules) realises a subset of the features associated with the verb form:

(3) Realisation rules for the Portuguese verb form *bebíamos* {Tense: Imperfect, Person: 1; Number: Pl}

a. $RR_{I, \{Tense: Imperf\}, V[Class2]}(X, \sigma) =_{def} \langle X\textit{ía}, \sigma \rangle$

b. $RR_{II, \{Pers: 1, Numb: Pl\}, V}(X, \sigma) =_{def} \langle X\textit{mos}, \sigma \rangle$

To clarify the format of RRs in (3), we provide the general format in (4) on which the realisation rules in (3) are based[6].

(4) $RR_{n,\tau,C}(X, \sigma) =_{def} \langle Y, \sigma \rangle$

Starting from the left, each RR carries three indices: the index *n* identifies the particular block to which the rule belongs; the index τ indicates the set of features that the rule realises, and the index C identifies the class of lexemes or subclass of lexemes (i.e. verbs, nouns, etc. or 2nd conjugation verb, 3rd declension adjective etc.) to which the rule applies. Accordingly, we say that the rule in (3b) belongs to Rule Block II (cf. Chapter 6 for details), realising the features '1st person plural' on a verbal category.

Like PFs, RRs are formulated as functions, which map a form and a set of features to another form. In (3), 'X' signals the input to the realisation rule; by default the input is the root of the lexeme[7]. This variant may also signal more complex inputs, in particular if previous RRs have

---

[6] In (4), angled brackets <...> denote an ordered sequence of elements, for instance, a root and a set of features; cruly brackets {...} denote an unordered set of elements, for instance, the feature set itself; parentheses (...) enclose the argument(s) of a function, for instance, the PF or RR. The argument may itself be an ordered set, in which case we get RR (<X,Y>). (cf. Stump 2001:44)

[7] The assumption that suffixes attach to a root will be later revised (cf. Chapter 6). For now, I shall restrict myself to the theory proposed in Stump (2001).

applied to the root. In the case of multiple affixation, the output of one rule serves as the input to the following rule. The right hand side of the rule therefore specifies the result of applying a given rule to X.

Let us look again at the verb form *bebíamos* 'we drank' and let us illustrate the application of realisation rules. Let us first assume that the root of the lexeme BEBER 'drink' is *beb-* and that the rule realising the innermost affix takes place first, as is standard in realisational morphology. We therefore apply (3a) to *beb-* yielding *beb-ía-* which is not yet a fully formed verb. However, this underived verb form serves as input to the rule in (3b), yielding *beb-ía-mos*. The base to which a rule applies may constitute the root of the lexeme or it may constitute a more complex morphological form. Unlike with PFs, the input to RRs is not necessarily the root of a lexeme and the output is not necessarily a cell in that lexeme's paradigm[8].

2.1.2.2 Defining a PF in terms of RRs

To better understand the relation between PFs and RRs we must think of PFs as being defined in terms of RRs. The general PF in (5) shows that a given Portuguese verb form Y is derived through the sequential application of RR_I and RR_{II}. In (6), the PF defines the Portuguese verb form *bebíamos*: the first line in (6) defines the root-feature pair to which the PF applies; the second line represents an ordered sequence of RRs which derive the verb form through successive rule application, and the third line expresses the inflected form which constitutes the morphological cell. In other words, the PF therefore defines the first-person plural imperfect cell in the paradigm of BEBER as the result of applying rules RR_I and RR_{II} (cf. 3).

---

[8] A notational observation: note that the combination of X with an affix is represented in (4) as Y. Note also that the complete feature set which defines the whole word form is represented as σ; it appears both in the input as well as in the output of RRs as well as in PFs.

(5) PF (<X, σ>) $=_{def}$ RR_{II} (RR_{I} (X, σ>)) $=_{def}$ <Y, σ>

(6) PF (<*beb*-,{Tense:Imperfect, Person: 1, Number: Pl}>)

= (RR_{II} (RR_{I} (<*beb*-,σ>)))

= <bebíamos,σ>

One key feature, then, of PFM is that PFs and RRs are hierarchically organised in the sense that PFs comprise the realisation rules which spell out the language's morphosyntactic properties. Also, rule application is always successive, with each rule applying to the output of the previous rule.

2.1.2.3 Organisation and application of realisation rules

Let us now consider how RRs are organised and survey some of the principles determining the application of RRs, following Stump (2001).

A word's inflectional affixes are often required to appear in a fixed sequence. If we consider, for example, the partial inflectional paradigm of the Swahili verb 'want' in (7), we observe that there are three distinct affix positions: position I occupied by tense markers, position II occupied by person/number agreement and position III occupied by negative polarity prefixes (Stump 1992, 1993).

(7) Fragment of Past tense paradigm of Swahili (Stump 1993)

	Positive				Negative			
	II	I			III	II	I	
1sg	ni-	li-	taka	1sg		si-	ku-	taka
2sg	u-	li-	taka	2sg	ha-	u-	ku-	taka
3sg	a-	li-	taka	3sg	ha-	a-	ku-	taka
1pl	tu-	li-	taka	1pl	ha-	tu-	ku-	taka
2pl	m-	li-	taka	2pl	ha-	m-	ku-	taka
3pl	wa-	li-	taka	3pl	ha-	wa-	ku-	taka

One assumption in PFM is that RR are organised into ordered rule blocks and that the set of affixes occupying the same position are realised by the same rule block (cf. also Anderson 1992). The rule blocks in (8) provide some of the rules we need to generate the partial paradigm in (7): the rules in (8a) realise the tense features, the rules in (b) realise subject agreement, and the rules in (8c) realise the feature polarity (the list of rules is not exhaustive; for full details, see Stump1992, 1993).

(8) Fragment of Realisation Rules for the Swahili verbal paradigm

a. Block I

i. $RR_{I\ \{Tense:Past\},\ V}$ (<X, σ>) $=_{def}$ <*li*-X,σ>

ii. $RR_{I\ \{Tense:Past,\ Pol:Neg\},\ V}$ (<X, σ>) $=_{def}$ <*ku*-X,σ>

b. Block II

i. $RR_{II,\ \{Per:2;\ Num:sg\},\ V}$ (<X, σ>) $=_{def}$ <*u*-X,σ>

ii. $RR_{II,\ \{Per:3;\ Num:sg\},\ V}$ (<X, σ>) $=_{def}$ <*a*-X,σ>

iii. $RR_{II,\ \{Per:1;\ Num:pl\},\ V}$ (<X, σ>) $=_{def}$ <*tu*-X,σ>

iv. $RR_{II,\ \{Per:2;\ Num:pl\},\ V}$ (<X, σ>) $=_{def}$ <*m*-X,σ>

v. $RR_{II,\ \{Per:3;\ Num:pl\},\ V}$ (<X, σ>) $=_{def}$ <*wa*-X,σ>

c. Block III

i. $RR_{III,\ \{Pol:Neg\},\ V}$ (<X, σ>) $=_{def}$ <*ha*-X,σ>

One further assumption is that the order in which rule blocks apply is determined by the rule block order, such that the affix closer to the root applies first. The order in which rule blocks apply (i.e., the template) is also specified by the PF itself, as the PF for the Portuguese verb form in (5), repeated in (6), shows[9]. For the Swahili verbal paradigm, we can provide the simplified PF given in

[9] In most cases, the PF defines the expected rule block order. Stump however contends that the sequence is not always invariant and that pardigm functions make it possible to account for more complex phenomena, particularly in those cases in which the default order of rule block application is violated, such as when position classes appear in reversed order (also known as affix metathesis). In Chapter 6, I shall refer to the phenomenon of affix reordering and it can be accouted for within PFM.

(9), which derives the 3-position template illustrated in (7), showing that all inflectional paradigms are 'templatic', since they always involve paradigm functions. In addition, each rule block can apply only once during the derivation of a word.

(9) PF for fragment of Swahili verbal paradigm (cf. Stump 1993)

$PF(\langle X, \sigma\rangle) =_{def} RR_{III}(RR_{II}(RR_{I}(X, \sigma\rangle)) =_{def} \langle Y, \sigma\rangle$

The affixes contained within a rule block are mutually exclusive give that they compete for the same position. In other words, the rules are disjunctively ordered, by virtue of Panini's principle, according to which only one rule will meet the required feature description, because only the more specific rule applies. In some cases, it is quite simple to figure out which rule, within a given rule block, applies in a given derivation, specially when there is only one rule that meets the required feature specification (the features {Person:1, Number: Pl} can only be realised by the rule in 8b-v).

In certain cases, however, more than one rule may seem to satisfy a given description. Consider, for example, the feature set {Pol:Neg, Tense:Past, Person:1; Numb: Pl} associated with TAKA, meaning 'we did not want'. The problem we face is to determine which one of the rules in block I derives the tense marker. Is it (8a-i) or (8a-ii)? Both are applicable, in that they tell us to add a prefix to realise the past tense feature. This type of rule block competion is resolved in PFM by Panini's principle. The principle says that when two rules could equally apply to a given form, the more specific rule applies in preference to the more general. This principle is often known as the Elsewhere Condition (Stump 2001 relables this 'Paninian Determinism'). In our example, the feature set {Tense: Past} is clearly more general than {Tense:Past, Pol:Neg}, so rule (8a-ii) applies rather than (8a-i) whenever we are dealing with a lexeme marked {Tense:Past, Pol:Neg}. When we come to a past tense lexeme which is marked {Pol: Pos}, only the more general rule (8a-i) can apply.

A second type of phenomenon which is common in inflectional morphology occurs when there is no formal marker for a given feature set. Take, for example, the future and past positive verb forms in (7) where an empty 'polarity' slot exists (i.e., affixal slot III). Morphemic models of morphology ought to postulate a zero morph realising the feature {Pol:Pos}. In PFM, on the contrary, all we need to say is that positive polarity is realised without applying any realisation rule. The key idea is that not being marked overtly can be just as significant as being marked with some sort of affix.

To capture this insight we will assume that every grammar contains one very general inflectional rule, known as the Identity Function Default (IFD). By this principle, if there is no explicit rule for realising a combination of features then we simply don't add anything. In other words, IFD applies to an argument and gives that same argument as its value: f(X)=X. In PFM, this is a universal default, applying within any rule block in case no explicit rule has applied. On this assumption, rule block III in (8c) yields the desired verb form without 'adding' any affix expressing {Pol:Pos}.

Summing up so far, central to the theory of Paradigm-Function Morphology is the idea that a language's inflectional morphology is defined through its Paradigm Function (PF). The PF (a theoretical construct unique to this theory) is a function from the root of a lexeme to one of the fully inflected words in the paradigm of that lexeme. In the morphology of any given language, a paradigm function is defined in terms of more specific realisation rules, for inflected words are often derived by applying two or more realisation rules. Realisation Rules (RRs) give morphological expression to a specified set of morphosyntactic properties. In this sense, we say that PFs and RRs are organised hierarchically.

One further key feature of the theory is that a lexeme's association with a particular set of morphosyntactic properties licenses the application of rules determining that lexeme's inflectional form. A language's realisation rules are organized into blocks; competition among realisation rules

belonging to the same block is resolved in favour of the narrowest applicable rule (Paninian Determinism Hypothesis). Every rule block also realises the Identity Function Default for the realisation of empty slots in the template. In addition, the short survey presented above has shown that Paradigm Functions enable Stump to capture generalisations which remain elusive in other theories of morphology (see Stump 2001, for full details).

2.1.4 Cliticisation and Paradigm Function Morphology

Spencer (2000) offers the first treatment of clitic systems within PFM, showing that RRs and PFs can be extended to account for the clitic system in Bulgarian and Macedonian. For Macedonian, for example, pronominal clitics are assumed to be the expression of featural content, as illustrated in (10) where the clitic *mi-*, for example, is derived as an exponent of '1st person singular dative' and the clitic *te-* as an exponent of '2nd person singular accusative'[10].

(10) Fragment of RRs for Macedonian clitics (Spencer 2000)

a. $RR_{III\ [CASE:\ Dat,\ PERS:\ 1,\ Num:\ Sg]}(X) \Rightarrow$ */mi-*X/

b. $RR_{IV\ [CASE:\ Acc,\ PERS:\ 2,\ Num:\ Sg]}(X) \Rightarrow$ */te-*X/

The crucial idea underlying the representation in (10) is that Macedonian clitics are not syntactically autonomous units, but instead combined morphologically with the verb. Similarly, it is assumed that object features (such as Case, Person and Number) are part of the range of features expressed by the Macedonian verbal paradigm. The representation of object clitics as featural content is also made in Stump (1980) and Miller&Sag (1997) for French, and Monachesi (1999) for Italian.

Minor extensions have been introduced to the theory of PFM to capture the attachment of clitic pronouns, in particular because clitics generally select hosts that are fully inflected. In the case of

[10] The notation of PFM in Spencer (2000) is a simplified version of Stump (2001).

Macedonian, the base to which clitic pronouns attach is an inflected verb form. The modifications however do not violate the spirit of the theory.

Spencer's (2000) insights will be further explored in thesis, following the proposals in Luís (2001a, 2003a, to appear) and Luís&Spencer (in press), by examining in detail various aspects about clitic placement and cluster formation in European Portuguese, in particular, and Romance, in general. Evidence is provided which motivates the need for a revised format of PFs and RRs. Again, the extensions do not change the formal foundations of the theory; instead they aim at capturing the similarities between inflectional and clitic systems. The main goal will be to look in detail at the idiosyncrasies of the clitic system in EP and to investigate how these properties can be insightfully accounted for within inflectional morphology. A preliminary analysis is given in Chapters 6 and 7; Chapter 9 offers a revised analysis; a summary of the proposal is given in Chapter 11[11].

The exact details of the analysis shall be discussed in the relevant chapters. Anticipating the details somewhat, it will be shown that a) RRs can derive clitic pronouns as verbal affixes, b) rule block ordering can be used to specify clitic order inside the clitic sequence and c) cliticised verb forms comprise two layers of affixation (i.e., the inner layer composed of ordinary affixes and the outer layer made of affixal clitics).

2.2 The grammar

In the study of the morphology of Europena Portuguese, relevant syntactic and phonological properties displayed by EP cliticisation made it necessary to examine the interaction between morphology and other components of grammar (in particular, syntax and phonology).

In our analysis of these interactions, an architecture of grammar is adopted which combines the realisational model of Paradigm-Function Morphology with Lexical-Functional Grammar

(Kaplan&Bresnan 1982), building on insights by Sadler&Spencer (2001), Luís&Sadler (2003), Sadler&Nordlinger (2003). LFG constitutes a grammar of parallel, co-present representations, namely Lexical-Functional Grammar (Kaplan&Bresnan 1982, Bresnan 2001, Dalrymple 2001). Fundamental to the LFG architecture is the assumption that each module of grammar has a distinct formal character and models a different aspect of the structure of language. Within this parallel design of grammar, I assume that words are derived in the morphological component and that the rules and principles determining a word's inflectional form are properly morphological, distinct from those of the syntax or the phonology (cf. also Stump 2001 for similar view of grammar). Thus, the output of the morphological component are the fully inflected words which constitute the terminal elements of the syntax, being inserted as leaves on syntactic trees. The separation between syntax and morphology has been stated as the lexical integrity hypothesis in Chomsky (1970), Lapointe (1980), Bresnan (2001).

Syntactically, LFG distinguishes between (external) phrase-structure from (internal) grammatical functions. Structural relationships are represented as c(onstituent, categorial)-structure, while grammatical functions are represented as f(unctional, feature)-structure. The separation between these two levels enables us to capture the ability for the same functional information to be realised by either phrases or parts of words, and also the fact that arguments of the verb are not bound to specific argument positions in phrase-structure. Furthermore, LFG syntax is non-derivational; the structural position of cliticised verbs will therefore be discussed from the point of view of the surface phrase structure (cf. chapter 10).

Among the implication of placing realisational morphology in correspondence with LFG-syntax is the fact that affixes can no longer be viewed as sub-lexical entries. So whereas in classical LFG, affixes (as sub-lexical entries) are associated directly with f-structure information, in 'realisational'

---

[11] Recent developments of the theory of PFM are also presented in Sadler&Nordlinger (2002), Spencer (2003, to appear), Otoguro (2003), Popova (to appear), Ackerman & Stump (to appear).

LFG an explicit correspondence between f-structure information and morphological features must be assumed to account for the interface between morphological structures and the syntax (Sadler&Spencer 2001, Sadler&Nordlinger 2002, Kaplan&Butt 2002, among others). I return to the morphology-syntax mapping in Chapter 10.

At the phonological level, I assume that phonological effects that are either irregular or conditioned by the identity of particular affixes or word categories take place within the domain of the morphological word being therefore instances of morphophonology. Phrasal phonology, on the other hand, applies productively across word boundaries (Stump 2001, Anderson 1992, Aronoff 1976, 1992). The morphophonology of clitics is discussed in various stages of the thesis: chapter 4 presents clitic-induced allomorphy and chapter 7 sketches an analysis; chapter 5 in the oother hand is devoted to phrasal phonology and examines previous studies within Prosodic Phonology (Nespor&Vogel 1986).

2.3 Summary

This chapter laid out the relevant parts of the framework within which EP cliticisation shall be analysed. I assume that Lexical-Functional Grammar (Kaplan&Bresnan 1982, Bresnan 2001, Dalrymple 2001) provides the overall architecture within which the realisational theory of Paradigm Function Morphology, as an autonomous module, interacts with other components of grammar. A more detailed discussion of PFM can be found in Stump (2001) and Spencer (ms). Detailed studies about the LFG architecture include Bresnan (2001), Dalrymple (2001) and Falk (2001).

Chapter 3 Clitics as morphology

This chapter surveys some of the morphological claims made about clitic pronouns in EP and other Romance languages. Section 3.1 discusses two prevalent morphological views of cliticisation. Drawing on data from French, Spanish and Italian, it is shown that both the inflectional view (Auger 1994, Miller&Sag 1997, Monachesi 1999, Brines 2001) and the morpho-lexical approach (Simpson&Withgott 1986, Bonet 1991, Grimshaw 1997 and Gerlach 2001a) argue against purely syntactic treatments, as initiated by Kayne (1975, 1991). Both types of approach, the inflectional and the morpho-lexical, however make substantially different claims about the exact morphological status of clitic pronouns: while the inflectional view regards clitic clusters as sequences of verbal affixes, deriving cliticisation as an inflectional phenomenon, the morpho-lexical approach regards clitics as lexical units with the ability to combine into compound-like clusters. Thus, morphology plays different roles in the placement of clitics in each one of the views. In section 3.2, I shall first examine how this debate extend to European Portuguese, and shall then formulate the position adopted in this thesis about to the role of morphology in cliticisation.

3.1 Approaches to cliticisation

3.1.1 Introduction

Two broad categories of clitics are usually referred to in the literature: phonological and syntactic clitics. Phonological clitics, also known as *simple clitics* (Zwicky 1977), are defined as prosodically deficient function words: they are unable to bear lexical accent and must attach to a stressed 'host' in phrasal phonology (cf. Selkirk 1995 on English; Booij 1995, Berendsen 1986 on Dutch; Vigário

1999a, Leeuw 1997 on EP). Since they are normally related to a corresponding full form with a similar distribution, their grammar is generally unexceptional[12]. A typical example of this type of clitic is the weak form of the English conjugation *and*:

(1) Bridget *and* John are waiting ([ɵn], [n̩])

Syntactic clitics, also known as special clitics are also regarded as word-level units. Unlike simple clitics however they are notorious for their special distribution in the clause given that they appear in positions generally not available for words with similar function (Zwicky 1977). They are frequently referred to as syntactic clitics and most of them share with simple clitics the property of being prosodically deficient.

Depending on the position they occupy in the clause, it has been argued that there are different types of so-called special clitics (Halpern 1999). Wackernagel clitics, for example, select the second position in the clause, i.e., after the first word or after the first phrase, such as Tagalog adverbials (Klavans 1980, Anderson 1993) or Serbo-Croatian auxiliaries and pronouns (Halpern 1995). Special clitics can also be of the Tobler-Mussafia type because they appear generally preverbally, except if that would make them 'sentence-initial', in which case they appear postverbally. These are found in Old French and Bulgarian (Hauge 1972). Other types of clitics are found in West Flemish where weak pronouns are third in V2 main clauses, but second in embedded clauses and V1 main clauses (Haegeman 1992).

(2) a. Taj *je* covek voleo Mariju. (Halpern 1995)
that CL man love.ppl Maria
'That man loved Maria'

b. Taj covek *je* voleo Mariju.

[12] Some so-called simple clitics (also known as phonologically weak function words) display rather complex morphological behaviour. For a discussion about English reduced auxiliaries see Luís (1997), Sadler (1997b), Barron (1997), Westcoast (2002).

Based on Zwicky's (1977) *simple clitic* vs. *special clitic* typology, some syntacticians have taken the position that pronominal clitics in Romance should also be regarded as special clitics, that is as word-level units with special syntax. The special distribution of Romance clitics, which is quite distinct from that of their full form pronouns, appears to be one of their most striking distributional properties (Kayne 1975, Cardinaletti&Starke 1999). For French, Kayne (1975), observed that unlike full pronouns pronominal clitics fail to a) appear in isolation (3), to be focused (4), to be coordinated (5) and to be modified (6):

(3) Qui as -tu vu? Lui/*Le.
who have -you seen? him
'Who have you seen? Him.'

(4) C'est lui / *le que j'ai vu.
it-is him that I have seen

(5) J'ai vu lui /*le et Marie.
I-have seen him and Marie

(6) J'ai vu lui /*le seul.
I-have seen him alone

Such evidence has motivated the view, initiated by Kayne (1975), that pronominal clitics should be treated as syntactically autonomous units subject to clitic-specific movement operations. In particular, clitics are assumed to be generated in argument position as the head of a complement DP and then to move as an Xº into their surface position. Kayne's proposal is also known in the literature as the 'movement' approach to cliticisation and has been widely adopted within derivational syntax for a wide number of Romance languages (e.g., Belletti 1999, Kayne 1991, Roberts&Cardinaletti 1991, Madeira 1993) [13].

[13] A detailed survey of syntactic treatments of cliticisation can be found in Riemsdijk (1995).

The view that clitics should be represented on syntactic nodes is not restricted to derivational theories of syntax but has also been proposed by Grimshaw (1982) and Schwarze (2001) within a lexicalist theory of syntax (i.e., Lexical-Functional Grammar, Kaplan&Bresnan 1982). In particular, Grimshaw (1982) assigns a syntactic position to each clitic, making use of phrase structure rules to determine the order in which they combine in the clitic string; Schwarze (2001), also places clitics in the syntax, but suggests that clitic clusters constitute lexically listed portmanteau units. There is a crucial difference between Grimshaw's early LFG-analysis of French and Schwarze's (2001) more recent analysis of Italian; in particular, the non-productive aspects about cliticisation, such as the structure of clitic sequences, is accounted for in the lexicon/morphology in Schwarze's analysis. In this respect, his work also adopts the view that clitic clusters pose problems to purely syntactic derivations, a claim put forward by Perlmutter (1970) and since then supported by a wide number of linguists.

In the next section, I will start by surveying two lines of analyses which argue that clitic pronouns in Romance exhibit morphological behaviour. In section 3.1.2, I summarise Gerlach (2001a), who assumes that clitic clusters should be derived as morphologically complex units and placed in the syntax (cf. Schwarze 2001). A quite distinct type of morphological approach is considered in section 3.1.3 with an overview of studies by Monachesi (1999), Miller (1992), Miller&Sag (1992), Auger (1994), Brines (2001). These latter studies, regard clitic pronouns in languages such as French, Spanish and Italian as verbal affixes and derive cliticised verb as inflected word forms, unlike Gerlach (2001a). In section 3.1.4, it is argued that an inflectional approach offers a more insightful and parsimonious treatment of the morphology of cliticisation.

3.1.2 Between the syntax and the lexicon

Studies by Simpson&Withgott (1986), Bonet (1991), Grimshaw (1997), Gerlach (1997) and Schwarze (1999) have focused mostly on the properties of clitic clusters and argued, based on the

underlying assumption that the rules of syntax are maximally general, that clitic sequences should be analysed outside the domain of syntax. In what follows, I will survey Gerlach's arguments against a non-syntactic derivation of clitic clusters and summarise her treatment of the data.

Gerlach's arguments against linearising clitics in the syntax are based mostly on the wide number of co-occurrence restrictions found within clitic clusters (cf. early generative studies by Perlmutter 1970 and Emonds 1975). For example, in most of the Romance languages, the combination between between dative and accusative clitics is determined by person features (e.g., 1st and 2nd person clitics can never combine with each other (7a), although they can combine with 3rd person clitics). Also, accusatives must generally follow datives (7b), even though the usual order for full NPs is in the exactly opposite order (7c).

(7)	a.	*me	te	→	*te	mi	(It., Gerlach 2001a)
		dat.1sg	-acc.2sg		dat.2sg	-acc.1sg	
	b.	*lo	mi	→	me	lo	
		acc.1sg	-dat.2sg.	→	dat.2sg	-acc.1sg	
	c.	Martina	spedisce	la	lettera	a	Vito.
		'Martina	sends	the	letter	to	Vito'

Given that full nominal phrases are not subject to these restrictions, Gerlach assumes that these co-occurrence restrictions cannot be argued to follow from general syntactic principles (cf. Chapter 4 for discussion). Further supporting this view is Bonet (1991) who observes that dialectal variation in clitic ordering inside the clusters in the Romance languages does not correlate with any syntactic or semantic differences. Based on this evidence, Gerlach (2001a) develops a meticulous analysis of clitic clusters and makes an important contribution to the view that cluster formation cannot be derived syntactically. Withihn Optimality-Theory, she introduces morphological alignment constraints which derive clitic sequences as combinations of two discrete clitic-units; cluster-specific idiosyncrasies, such as rigid ordering, co-occurrence restrictions and non-productive shape

variation, follow from other clitic-specific constraints (Grimshaw (1997) and Bonet (1991) also regard clusters as morphologically complex units). As to cluster placement, it is argued that the syntax provides clitic-specific placement rules which attach clitics to Vº, both preverbally or postverbally.

Even though Gerlach's analysis correctly argues that clitics do not constitute word-like units and that, consequently, clitic clusters do not behave like word sequences, her analysis also argues forcefully that clitics do not behave like affixes. The exact details of Gerlach's analysis shall not concern us in this chapter (cf. Chapter 7), but it is worth observing that Gerlach (2001a)'s claims entails the view that clitics constitute entirely distinct categories in grammar that are listed in the lexicon as [-affix, -word] units (vs. words [-affix, +word] and affixes [+affix, -word]). Gerlach's analysis is also 'costly' from a conceptual point of view because she introduces clitic-specific operations to handle the phonological, lexical and syntactic aspects of the data: a) clusters are derived through clitic-specific lexical constraints; b) cluster placement through clitic-specific syntax, and c) clitic induced shape variation through clitic-specific phonology.

The crucial for the present thesis is whether clitic pronouns are as distinct from affixes as Gerlach's analysis assumes. If it can be shown that clitics are indeed too exceptional to be assimilated to an available category, such as the affix, then we will agree that clitic-specific mechanisms may be needed to capture the data. However, what I will try to show next is that the behaviour of clitic pronouns in Romance resemble very much the behaviour of affixes and that Gerlach's analysis fails to account for these clear similarities.

3.1.3 Inside inflectional morphology

The claim that clitic sequences should be analysed outside the domain of syntax is also argued for by inflectional approaches to cliticisation (cf. Chapter 9). However, unlike the work of Gerlach, the non-syntactic nature of cliticisation is captured by treating clitics as verbal affixes and by deriving

cliticised verbs as inflected words (Miller 1992, Miller&Sag 1997, Monachesi 1999, Brines 2001). Crucial evidence in favour of this position is the wide range of idiosyncratic phenomena (cf. chapter 4) which falls outside the scope of syntactic operations (Zwicky&Pullum 1983, Zwicky 1985, 1987, Smirniotopoulos&Joseph 1998).

Among the data supporting the affixal status of clitics is, for example, the immediate adjacency between clitics and the verb; as is well known , clitics in French, Italian and Spanish cannot be separated from the verb. One further piece of evidence is provided by the fact that clitics only select hosts belonging to the same category, namely the category of verbs. In addition, there is the fact that clitics and verbs undergo non-productive allomorphy: this means that the phonological changes suffered by clitics or by the adjacent verb do not occur across word boundaries but are instead determined by grammatical features either of the verb or of the clitic, or both. Further evidence supporting the affixal nature of clitics is provided by clitic clusters. Inflectional studies also assume that cluster-internal idiosyncrasies do not follow from syntactic principles, regarding the non-syntactic properties (i.e, rigid ordering, idiosyncratic co-occurrence restrictions and shape variation) as a consequence of the affixal status of clitics. Rigid clitic ordering is therefore argued to resemble the rigid order of affixes (Anderson 1992, Cummins&Roberge 1994, Auger 1994, Monachesi 1999, Miller&Sag 1997, among others).

Based on these and other arguments (cf. Chapter 4), inflectional studies have been developed for a large number of Romance languages, within different theoretical models (e.g., Auger's 1994 A-Morphous Morphology account; the HPSG analyses by Miller&Sag 1997, Monachesi 1999, Brines 2001, Crysmann 2002; Nishida's (1987) Categorial Grammar approach, among other). Regardless of which formalism is adopted, these authors analyse the clitic system as a fundamentally inflectional system (c. Chapter 6 and 9 for further details). To conclude, then, inflectional studies capture the non-syntactic behaviour of clitics without introducing an additional

category into the lexicon. Clitics are assimilated to the category of affixes (Monachesi 1999) and inflectional mechanisms are adopted for the derivation of cliticised verbs (cf.Chapter 6).

3.1.4 Discussion

In what follows, I will show that there are both empirical and conceptual reasons for preferring an inflectional analysis of cliticisation.

a) the properties of the verbal host

One of Gerlach's (2001) claims against an inflectional approach is based on the idea that clitics are less selective than ordinary affixes; in other words, clitics select inflected verb forms while affixes select specific inflectional stems. The example provided draws on Italian subjunctive forms, as given in (8), where the 1st and 3rd singular agreement suffix *-i* selects first conjugation stem, while the suffix marker *-a* selects second and third conjugation stems. This type of morphological conditioning of verbal suffixes is prevalent throughout Romance.

(8) 1sg/3sg cant-i ced-a finisc-a (Gerlach 2001a:46)

From an inflectional point of view, it may however be argued that even though clitics don't select the conjugation class of the host, they are nonetless selective with respect to the host they attach to (Miller&Sag 1997; Brines 2001, Monachesi 1999). As alluded to before, clitics select the category of the host (i.e., the verb) and are in various ways sensitive to the verb's morphological and morphophonological properties. In French, for example, clitics are sensitive to the tense features of the verb: the locative clitic *y*, which is subcategorised for by the verb *aller*, cannot combine with all of its the future forms. While it can combine with other forms of the verb, such as *vais* or *vas*, it is prevented from occurring with *irais* '(I) will go', as in as in **j'y irais*. It may be argued that this restriction is motivated by purely phonological factors, however it has been shown that adjacent

sequences of [i] are not prevented from occurring elsewhere (Miller 1992, Auger 1994, and references therein).

There are also morphophonological effects found at the verb-clitic boundary. These are generally associated with specific grammatical properties of the verb and suggest that clitics are sensitive to the properties of the host. In French, for example, liaison of the nasal consonant [n] is only possible word-internally between affixes and their stems; however as (9) shows it also takes place between clitics and the verbal host (Miller 1992, Miller&Sag 1997).

(9) a. Marie en a. ([ãna] / *[ãa])
Marie cl has
'Marie has some'

b. ira-t-on à la ville? ([iratõalavil] / *[iratõnalavil])
go-t-cl to the town
'Are we going into town?'

In addition, we also find an idiosyncratic realisation of the 1st singular present tense form of *être* when combined with the subject clitic *je*. In this case, *je suis* is realised as *chuis* [šɥi], but the same does not apply to the homophonous sequence *je suis* meaning 'I follow' (cf. also Auger 1994 and references therein, for Québec Colloquial French). Other examples include the elision of [y] in *tu* before vowel-initial verbs as in *t'es* 'you-sg are' which is not found elsewhere in the language.

Further examples of shape variation at the verb-clitic boundary are found in colloquial Castilian Spanish (Brines 2001, Menendez Pidal 1904, Fernández Soriano 1993). There is consonant assimilation when consonant-final verbs are followed by a clitic (10); consonant metathesis takes place on imperatives that are followed by a consonant-initial clitic (11), and imperatives that are followed by a 2nd person plural clitic pronoun can either undergo verb-final consonant deletion (12a) or the infinitive form of the verb may be used instead (12b):

(10)	a.	*servirle	→	serville	'To serve him'	(Brines 2001)
	b.	tornarse	→	tornase	'To become'	
(11)	a.	dadle	→	dalde	'Give him/her!'	
	b.	dadnos	→	dandos	'Give us!'	
(12)	a.	*Callados	→	Callaos.	'Shut up!'	
	b.	*Callados	→	Callaros.	'Shut up!'	

These non-productive phonological alternations are only found at the verb-clitic boundary. Under a Gerlach-style of analysis, they would have to be accounted for through clitic-specific rules of phrasal phonology, such as vowel or consonant truncation in edge positions. To support such view, it has been argued that the type of phonological variation triggered by clitics on the verb is also found across word-boundaries and that therefore no empirical motivation exists for an affixal analysis of clitics. This view has for example been adopted for EP by Vigário (1999b, cf. chapter 5). However, strong evidence suggests precisely the opposite, namely that these shape alternations do not take place between other adjacent word units (see Chapter 7 for discussion).

Other morphophonological changes are found within the clitic cluster. For example, in Italian, the clitics *mi, ti, ci, vi* are replaced by *me, te, ce, ve* immediately before *ne, lo, la, li, le*, as in (13a). Similarly, clitic substitution is found in Spanish whenever the dative object clitic *le* is followed by the accusative clitic *lo, le* is replaced by the reflexive clitic *se*, as in (14) (cf. also Bonet 1991, for Catalan). What the data illustrates is that these alternations are triggered by phonological properties of adjacent clitics.

(13)	a.	*ci	la	→	ce la	(It.)
		loc	acc.3sg.fem			
	b.	ci	si	→	*ce si	
		loc	refl.3			
(14)	*le		lo	→	se	lo (Sp.)
	dat.3sg		acc.3sg.masc		refl.3sg	acc.3sg.masc

In 'ordinary inflection', phenomena like these are very common. For example, the Finnish plural marker is realised as [t] in the nominative form, but realised as [i] if it is followed by a case marker (e.g., *talo-i-ssa* 'in houses' - Plural Inessive form of *talo* house; *talo-t* 'houses' – Plural Nominative form of *talo*). European Portuguese clitics, as we shall see in Chapter 4, undergo a wide range of non-productive alternations, supporting the claim that clitic pronouns behave more like affixes.

It might be argued that these cluster-internal effects could be derived through clitic-specific phonological rules. The problem, however, would be to account for the similarity between allomorphic effects found with ordinary affixes and with clitics. Under an inflectional analysis, this similarity falls out naturally.

b) distribution of clitics

Non-syntactic effects are also reflected in the distribution of clitics. Even though Gerlach assumes that cluster placement is determined by syntactic operations, it is far from obvious that this approach really offers an insightful account. Note that in Spanish and Italian clitic pronouns only ever appear immediately before or after the verb, suggesting that the clitic position is ultimately determined by the need to select the category of the host rather than by the need to appear in a given syntactic position (Nishida 1987, Brines 2001, Monachesi 1999, Miller 1992, Miller&Sag 1997). Indeed, Gerlach (2001a) draws a similar conclusion and argues – on the basis of the high degree of locality – that clusters should attach syntactically to a V° node. The problem, however, seems to be the zero-level host which is arguably syntactically selected. If indeed, clitics display such as high degree of selectivity and of locality with the host, then it seems rather strange to treat them as word-level units. It would be more elegant and more simple to simply assume that clitics are in fact sensitive to the category of the host, like affixes; the fact that the host is V° and not another X-bar category also

indicates that the clitic and the host form a cohering unit which cannot be separated. This, again, resembles the morphological dependency of affixes. In fact, the similarity between clitics and affixes is further supported by the fact that clitics in French, Spanish and Italian cannot be separated from the verbal host. The inseparability follows without any further stipulation under an inflectional analysis. But, if clitics are placed in the syntax, as suggested by Gerlach, then the cohering relation appears to be not well motivated.

Even the fact that Spanish clitics appear before and after the verb cannot be regarded as being typical of function words (compare the place of determiners with respect to nouns). On the contrary, the ability for a given ordinary affix to appear either as a suffix or as a prefix has been attested as an inflectional phenomenon by Noyer (1994) and Stump (1993). The fact that there are languages in which ordinary affixes exhibit some mobility indicates that the 'dual' placement of clitics should not be used as evidence in favour of a syntactic analysis. There are however more arguments suggesting that an inflectional analysis offers the most insightful account. As mentioned above, there is the local relation between clitic pronouns and the verb, and then there is the arbitrariness of the factors determining preverbal and postverbal placement. As observed by Cummins&Roberge (1994), features like finiteness or imperative can have different effects on clitic placement proclisis and enclisis have different effects in French and Italian varieties, , further supporting the idea that placement is arbitrary and cannot be motivated in terms of deep syntactic properties.

3.1.5 Summary

This section considered two morphological approaches to cliticisation in Romance. It was argued that they differ significantly with respect to the role they assign to the morphology in the derivation and placement of clitics: a) inflectional analyses, as proposed by Miller (1992), Miller&Sag (1997), Monachesi (1999), among others (cf. 3.1.3), assimilate clitics to well-established affixal categories; b) on the contrary, the analysis developed by Gerlach (2001a), following Bonet (1991) and

Grimshaw (1997), among others, prefers to adopt a partly lexical and partly syntactic approach (3.1.2). Under the latter view, clitics are assumed to be distinct from affixes and constitute distinct theoretical categories with their own set of phonological, syntactic and morpho-lexical properties. An appraisal of the inflectional approach was made showing that there are clear similarities between affixes and clitic pronouns which are more naturally accounted for under an inflectional analysis.

3.2 Portuguese clitic pronouns

This section addresses some of the literature on the clitic system of European Portuguese (EP). This language forms the empirical basis of this study and will be analysed in the fortcoming chapters from an inflectional point of view. The position taken about the EP clitic system is that it is inherently inflectional, with enclitics constituting genuine verbal suffixes and proclitics behaving like phrasal affixes (Luís 2003a, to appear). It is argued that the interaction between word-level affixation and phrase-level affixation is what uniquely characterises the clitic system of EP. First however, let us survey previous claims that have been made (for more detailed surveys, cf. Chapters 5-8)

3.2.1 Overview

EP shares with other Romance languages the fact that clitic pronouns are distributionally distinct from full pronouns: clitics cannot be used in isolation (15b), they cannot be focused (16b), they cannot be coordinated (17b), they cannot be modified (18b).

(15) a. Quem viu o jogador? Ele.
who saw the player? he
'Who saw the player? He did'
b. A quem deste o livro? *lhes

to whom gave the book? dat.3.pl
'Who did you give the book to? Them'

(16) a. Foi ele que me deu o livro.
was he who me gave the book
'It was him who gave me the book'

b. *Os polícias viram te e não me'
the policemen saw acc.2sg and not acc.1sg
'the policemen saw you but not me'

(17) a. Tu e eu fomos ao jogo.
you and I went to-the game
'You and I went to the game'

b. *O jogador viu o e a.
the player saw acc.3sg.masc and acc.3sg.fem
'The player saw him and her'

(18) a. Até ela telefonou.
even she called
'Even she called'

b. *Até te elas viram.
even acc.2sg they saw
'Even you, they saw'

In line with the tradition initiated by Kayne (1975, 1991), and continued by his followers, the evidence shown in (15-18) has indicated to some that EP clitics constitute special word-level units and that the verb-clitic combination should be generated through syntactic movement operations (see Duarte&Matos 2000 for survey of syntactic approaches of the EP data).

Despite this line of study, the behaviour of clitic pronouns in European Portuguese has recently generated a heated debate in the literature due to the growing assumption that cliticisation cannot be accounted for without taking into account the morphological nature of the data. Departing from strict syntactic approaches (Madeira 1993, Martins 1994), recent analyses have appeared within different theoretical frameworks arguing that purely syntactic accounts are not tenable (e.g., Vigário 1999a, Crysmann 2002, Gerlach 2001a, Stolz 1992, Zwicky 1987). In these accounts, the exact

extent to which morphology plays a role in the derivation of clitic structures varies quite significantly. For example, Vigário (1999b) proposes a phrasal allomorphy treatment within Precompiled Phonology, Gerlach (2001a) combines clitics in the lexicon, and Crysmann (2002) who develops a linearisation-based treatment within HPSG which generates clitics as 'hybrids'[14]. In this thesis, I also assume that purely syntactic treatments cannot account adequately for the data, however, unlike previous studies by Vigário (1999a), Crysmann (2002) and Gerlach (2001a), I adopt an inflectional view of cliticisation.

Among the data that will be presented in detail in Chapter 4, it is shown that EP clitics exhibit shape variation which is highly dependent on the grammatical features of both the verb and the clitic. For example, postverbal *nos* 'us' or *vos* 'you.pl' trigger consonant deletion on the verb-final *–s*. The idea that clitics are sensitive to the properties of the verb is also illustrated by mesoclisis. In this case, clitics selects a specific verbal stem, i.e., the infinitival stem, under specific morphological conditions, i.e, if the verb form is future or conditional. In addition, there are cluster-internal idiosyncrasies, such as portmanteau effects, fusion, syncretism, rigid ordering and co-occurrence restrictions, that are uniquely triggered by clitics and that are typical of affixes.

These striking similarities between clitics and affixes motivate the inflectional status of the EP clitic system (e.g., Stolz 1992, Zwicky 1987**,** Luís 2000, 2001b). However, as pointed out in Crysmann (1997), standard inflectional approaches to cliticsation as developed for French (Miller&Sag 1997) or Italian (Monachesi 1997) cannot be adopted in their entirety for EP. While postverbal clitics (i.e., enclitics) show all the signs of the morphological attachment found in Franch and Italian, proclitics exhibit an intriguingly phrasal behaviour with distributional and scopal properties that are not found in these other Romance languages (cf. Chapter 8 for more details).

[14] These studies are discussed in more detail throughout the thesis at relevant points of the argumentation.

(19a) shows that preverbal clitics do not have to be strictly adjacent to the verb. (19b) illustrates that they can have wide scope over two conjoined VPs (Martins 1994, Matos 2000).

(19) a. ... que o João lhe não revelou o segredo.
... that the J. dat.3.sg not revealed the secret
'… that J. has not revealed the secret to him/her'

b. Todos me escreveram postais e enviaram prendas.
all dat.2sg wrote cards and sent presents
'All wrote me cards and sent me presents'

These then are the facts that any account of EP clitic pronouns must account for:

(20) a) postverbal clitics form a morphologically cohering unit with the verb;
b) preverbal clitics undergo interpolation and can have wide scope;
c) clitics and clitic clusters in preverbal and postverbal position are exactly identical:

(21) a. bebi -o → não o bebi
drank -acc.2sg.masc not acc.2sg.masc drank
'I drank it' 'I didn't drink it'

b. apresentou -se --me → não se me apresentou
introduced -refl.3sg -dat.1.sg not refl.3sg dat.1.sg introduced
'She introduced herself to me' 'She didn't introduce herself to me'

c. davas -mo → não mo davas
give -dat.1.sg/acc.3sg.masc no dat.1.sg/acc.3sg.masc give
'You gave it to me' 'You didn't give it to me'

With respect to (21c), it should be pointed out that the absolute homophony between clitics and clusters in preverbal and postverbal position suggests that we are dealing with the same unit, despite the asymmetric behaviour summarised in (21a) and (21b).

This section has just observed that the scopal behaviour of proclitics and the ability of proclitics to be separated from the verb challenge the idea - assumed for clitics in French, Italian and Spanish -

that clitics are part of verbal morphology. The EP clitic system therefore occupies a quite unique position within the context of Romance. This means that previous inflectional analyses of Romance clitic pronouns (e.g., Miller&Sag 1997, Monachesi 1999, Auger 1994) would fail to account for EP in its entirety.

3.2.2 An inflectional proposal

The goal of this thesis will be to show that the phrasal behaviour of proclitics, illustrated briefly in (19), does not invalidate an inflectional analysis of the EP clitic system. My main claim will be that the clitic system is inherently inflectional because enclitics, which constitute the default case in EP, behave typically like inflectional affixes. They are strictly adjacent to the verb, they combine into clusters and trigger/undergo allomorphic alternations. In particular, postverbal clitics, which constitute the default case in EP (Duarte et al. 1995)[15], behave like verbal suffixes, thus supporting to the inflectional status of clitics. The data shall be illustrated in much more detail in Chapter 4, with priority being given to the case of enclitics.

The crucial question then is how to accommodate the asymmetry between enclitics and proclitics within an inflectional approach to cliticisation (cf.19). As pointed out before, the fact that proclitics can undergo interpolation and have wide scope over coordination means that they do not form a morphologically cohering unit with the verb unlike postverbal clitics. To account for the enclitic-proclitic asymmetry, I shall first assume that proclitics and enclitics constitute the same affixal unit. Given that proclitics are formally exactly identical to enclitics, it would be counter-intuitive to argue that proclitics are categorically distinct from enclitics. However, since proclitic affixes cannot

[15] The idea that enclisis is more productive than proclisis is supported by data from child language and also by adult speech: in particular, while children learn postverbal cliticisation at an earlier stage

combine morphologically with the verb, unlike morphological affixes, it will be argued that proclitics constitute instead phrasal affixes, that is, affixes attached to phrasal node (following insights from Klavans 1980, Anderson 1992, Legendre 2000a,b; Spencer 2000).

Summing up then, this thesis argues that EP clitics should be analysed as inflectional affixes. The view is supported by the cohering relation between the verb and the enclitic, and the fact that enclitics constitute the default form of cliticisation in EP. The asymmetry between proclitics and enclitics is derived as a difference in status between the preverbal and postverbal placement of the same affixal unit. While postverbal clitics attach to the verb like verbal suffixes, preverbal clitics attach within a syntactic domain as phrasal affixes.

3.2.2.1 Two types of affixation

In what follows, I will briefly clarify the distinction between morphological affixation and phrasal affixation. Both types play an important role in the inflectional analysis of EP clitics developed in this thesis.

a) Morphological affixes

The term 'morphological affix' is used here to refer to the more ordinary types of inflectional affix. These affixes are bound morphological units which need a morphological base to attach to. The attachment takes place inside the morphology, through ordinary inflectional mechanisms (Matthews 1972, 1991; Anderson 1992; Stump 2001). Within realisation models of inflectional morphology, for example, the attachment of affixes is modelled through inflectional rules realising affixes. These rules take as input an underived base (either a root or stem) which they associate with a given affix

---

of the acquisition process (Duarte et al. 1995), adults quite pervasively choose enclitics in contexts where proclitics should occur.

yielding an inflected word form. Morphological (or word-level) affixes are therefore part of the morphology of word. (Cf. Chapter 2 for detailed exposition of inflectional derivation of affixes within the theory of Paradigm-Function Morphology).

The claim that clitic pronouns in various Romance language (e.g, French, Italian, Spanish, Rumanian) have the status of morphologically attached inflectional affixes has been formulated by Monachesi 1999, Miller 1992, Miller&Sag 1997, Auger 1994, Nishida 1986, Brines 2001. Formally, this means that clitic sequences constitute sequences of affixes that attach to the verbal host through inflectional morphology (cf. Stump 1980, Miller&Sag 1997:585, for French; Auger 1994, for Canadian French; Spencer 2000, for Macedonian, Luís 2000, 2001a,b for EP; Brines 2001, Heap 2000, Nishida 1987, for Spanish, among other). The idea that cliticised verbs constitutes inflected word forms shall be examined and formalised in detail in Chapters 6, 7 and 9 of this thesis.

b) Clitics as phrasal affixes

While morphological affixes select stems that are members of a specific word class, phrasal affixes select a specific position in the syntax. In addition, phrasal affixes do not form a morphologically cohering unit. While morphological affixes combine with their host in the morphology, forming an inflected word with it, phrasal affixes only attach to the host in the syntax (Anderson 1992). Let us briefly illustrated this idea with data on Serbo-Croatian.

Several arguments seem to show that clitic sequences in Serbo-Croatian resemble sequences of affixes (Anderson 1996, in press, Spencer 1992, O'Connor 2002). There is, for example, the fact that clitics combine into rigidly order clitic sequences as in Romance. In particular, the question particle *li* precedes auxiliary and pronominal clitics and is followed by auxiliary and pronominal (dative and accusative) clitics. One curiosity about clitic order is the behaviour of the 3rd singular pronominal clitic *je* which appears always at the end of the clitic string, after dative and accusative clitics. Neither the clitic sequence nor the special behaviour of *je* find an obvious motivation in the

syntax. Further evidence is provided by grammatically conditioned allomorphic alternations inside the Serbo-Croatian cluster with the pronominal clitic *je* 'fem. sg. acc' which surfaces as *ju* when followed by another clitic *je*. At the verb-clitic boundary, it has also been observed that infinitive verb forms undergo non-productive shape variation when followed by clitic forms of the future auxiliary (cf. Spencer 1992). It has been point out that these idiosyncrasies would not be insightfully accounted for if clitics in this language were treated as lexical items and combined into clusters through syntactic constraints.

Despite these affixal properties, Serbo-Croatian clitics do not select the category of the host they attach to. So, in the examples given above, Serbo-Croatian clitic *je* attaches to a determiner in (2a) but to a noun in (2b). In Ngiyambaa (Klavans 1980), as shown below, second position enclitics attach to an adjective as in (22a) following the first word, and to a noun in (22b) following the first NP.

(22) a. nadhay =ndu guya dha -yi
tasty =2Nom fish eat -Past
'You ate a tasty fish'

b. nadhay guya =ndu dha -yi
tasty fish =2Nom eat -Past
'You ate a tasty fish'

Both languages share similar placement restrictions. Their clitics satisfy the second-position requirement, namely, they must appear either after the first word and after the first phrase (cf. Chapter 8). The concept of phrasal affixation therefore entails that certain affixes do not select the category of their host[16].

[16] It is largely on the basis of 'promiscuous' clitics (i.e., clitics which do not select the category of their host), that Klavans (1980, 1985) formulates her theory of phrasal affixes. Anderson (1992), on the contrary, uses the concept of phrasal affixation as a synonym for cliticisation. In this thesis, we

So, what characterises these clitics is their placement in specific syntactic positions. Which node phrasal affixes attach to can be determined on the basis of placement parameters, as the treatments by Klavans (1980, 1985) and Anderson (1992, 1996) have shown (cf. Chapter 8). Recent studies have adopted the concept of phrasal affixation within a variety of theoretical frameworks (Nevis 1985; Miller 1992; Spencer 2000; Legendre 2000, in press; Anderson 1996, in press; Luís 2003a, O'Connor 2000, Otoguro 2003, Börjars 2003). I shall return to phrasal affixation in Chapter 8 where I will survey in some detail the treatment given by Klavans (1980, 1985) and Anderson (1992, 1996, 2000) to phrasal affixation[17].

3.2.2.2 The clitic system of EP

In this study, I assume that the distinction between word-level affixes and phrasal affixes, following Klavans (1980), Anderson (1992, 1995, 1996, 2000), Spencer (2000), Legendre (1997, 2000), Börjars (2003). Based on this assumption, I shall argue that EP has a mixed clitic system in which enclitics behave like word-level suffixes and proclitics like phrasal affixes. Under this approach, enclitics and proclitics are not different in kind because they are generated as affixes; instead the difference between enclitics and proclitics is regarded as a difference in status between word-level and phrasal placement of the same affixal unit or affixal sequence (cf. Chapter 8-9).

---

follow Klavans in assuming that phrasal affixes are affixal clitics which do not select the category of the host. On the contrary, affixal clitics which appear systematically on the head of a given phrase are best viewed as genuine morphological affixes (Spencer 2000, Börjars 2003).

[17] Some authors use the term 'phrasal affixation' as a synonym of 'edge inflection' (for instance, Crysmann 2000, 2002, explicitly equates these terms). In line with Halpern (1995) and Spencer (2000), I shall reserve the term 'edge inflection' for phenomena where an affix appears at the edge of a phrase but undergoes (or triggers) morphophonological alternations typical of ordinary affixes. Edge inflections therefore are morphologically part of the host, even though they tend to appear at the edge phrase (cf. English possessive in Zwicky 1987, Bulgarian and Macedonian determiners in Halpern 1995).

At this point, it seems important to ask why the phrasal-affix approach should not be adopted for enclitics as well. If it was, then the analysis would appear more homogeneous. Of course, the problem of doing this is that the data, as shown above, is not symmetrical. So, deriving enclitics as phrasal affixes would fail to explain various crucial aspects of the data (cf. Chapter 4 for illustration): a) the fact that enclitics cannot be separated from the verb and can only ever occur after a verbal base, b) the abundance of allomorphic variation with enclitics and the complete absence of it with proclitics, and c) the ability for postverbal clitics to appear verb-internal after an infinitival stem. A phrasal affix analysis along these lines is in effect provided by Anderson (in press) for Romance clitics, including EP. I shall return to the problems raised by this approach in Chapter 8. For now, it suffices to say that the empirical facts observed for enclisis fall out more naturally under morphological affixation. Therefore, a mixed inflectional approach of EP cliticisation seems to offer a more insightful analysis than a uniformly phrasal-affix approach. In effect, enclitics and proclitics constitute different affixal categories.

To capture this insight, I will propose an analysis which assigns affixal clitics the ability to be attached either to a phrasal node or to be placed postverbally with respect to a verbal stem. The proposal assumes that a set of *exponence rules* which derive a clitic sequence, such as *se-me* '3rd.refl-1stdat', as composed sequences of affixes; it is further assumed that *placement rules* determine both the direction of attachment of a given clitic cluster and the type of host with respect to which the cluster is placed (cf. Chapter 9 for detailed analysis). The analysis therefore enables placement rules to refer either to morphological or to phrasal hosts, thus capturing the fact that clitics in EP appear on either type of host[18]. This insight will be explored and formalized within the theory of Paradigm-Function Morphology (Stump 1993, 2001), adopting insights from (Luís 2001a, 2003a, to appear), Luís&Spencer (in press), and Spencer (ms).

---

[18] It shall be argued later on that the 'mixed' attachment of inflectional clitics is not only attested for EP. In Chapter 10, I refer to the case of the Greek dialect of Cappadocia (Janse 2000).

3.2.3 Earlier accounts

The insight that EP clitics display morphological behaviour is formulated in Gerlach (2001a), Vigário (1999b) and Crysmann (1997, 2000). These studies make different claims about the role of 'morphology' but share the idea that there are non-prodctive effects in cliticisation which the syntax alone cannot capture. Gerlach's proposal has already been summarised in section 3.1, with respect to other Romance languages. I shall not addressed it here because the claims made about Romance also extend to EP (however cf. Chapter 7 for a discussion of Gerlach's treatment of phonological shape variation). I will instead refer to Vigário (1999b) and Crysmann (1997, 2000). Both studies shall also be discussed at various points of the argumentation throughout the thesis. Only a brief outline of their main claims is given here.

VIGÁRIO (1999b)

In this study, cliticisation is analysed as a prosodic process that combines unstressed word units (also known as function words) with a stressed host. Special emphasis is given to clitic shape variation and the phonological effect of postverbal clitics on the verb. Vigário is aware of the fact that clitic induced allomorphy cannot be adequately accounted for through pure phrasal phonology, she proposes an account within Hayes (1990)'s theory of Precompiled Phonology is proposed. The main argument is that clitic pronouns have phrasal allomorphs (i.e., allomorphs that are listed in the lexical entry and that can be selected in the syntax, cf. Chapter 7). In short, then, Vigário's work derives the verb-clitic unit as a syntactic combination, but adopts a 'precompilation' approach to explain why clitics induce and undergo non-productive shape variation in postverbal position.

I will now briefly point out some of the problems with Vigário's analysis. Her approach shall be discussed in more detail throughout the thesis (cf. in particular, Chapters 5 and 7). To begin with, her analysis derives enclitics as 'exceptional' function words and regards proclitics as the

'unmarked' units. Among her claims about enclitics is the idea tha they incorporate into the host instead of adjoining to it like ordinary function words (Vigário 1999a); and that they can be realised through phrasal allomorphs, unlike proclitics. However assigning enclitics 'marked' grammatical status is completely at odds with the well-known fact that postverbal clitics constitute the more productive and thus the unmarked realisation (cf. Duarte et al. 1995, Duarte&Matos 2000).

The idea that shape variation can be derived within Precompiled Phonology also seems quite artificial, particularly because clitic-induced allomorphy in EP is triggered by properties such as person and number features. Grammatically conditioned allormophy, as Hayes (1990) points out, belongs into the domain of inflectional morphology (cf. Chapter 5 for discussion). Also, in some instances of grammatically conditioned variation, the allomorphy is 'reciprocal', that is, it affects both the clitic and the context of insertion of the clitic, i.e., the adjacent verb (cf. Chapter 4). In Hayes (1990), however, once phrasal allomorphs are inserted they may have no effect on their own context.

Finally, there also seems to be insufficient empirical data to support the claim that pronominal clitics behave phonologically like function words. Furthermore, the rigid order inside clitic sequences and co-occurrence restrictions between clitics are analysed by treating clusters as opaque portmanteau units; clusters are thus listed in the lexicon side by side with individual clitics (cf. Chapter 9). The problem with this account of clitic clusters is that it is purely stipulative, failing to account for the transparent internal structure of clitic sequences.

In Chapter 5, some of Vigário's arguments will be re-examined and it is argued that the evidence in favour of viewing enclitics as word-level units is quite weak. I shall return to this analysis later in the thesis, for now I conclude without further arguments that Vigário's analysis, as a whole, is not tenable.

CRYSMANN (2000, 2002):

Within HPSG (Pollard&Sag 1992, Kathol 2000), Crysmann assumes that clitics in EP are generated as 'hybrid' morphosyntactic units, partly affixes and partly words, resembling separable verb-particles which can either attach to the verb as a kind of 'prefix' or be separated from it by whole syntactic phrases (cf. German or Hungarian, in Kathol 1995, Ackerman&Webelhuth 1998, Lüdeling 2001). Standard analyses of separable verb-particles argue that when particles appear as prefixes, attached to the verb, they form with it a morphological unit. Evidence supporting this claim comes from the fact that particle-verb combinations participate in morphological operations, e.g., derivational morphology (Ackerman&Webelhuth 1998). On the contrary, when the particle is separated from the verb by whole phrases (e.g, German: Sie *schaut* alle Männer *an*. 'She looks at any man'), the particle is analysed as a syntactic word (Lüdeling 2001, for German).

The analogy between clitics and preverbal particles is explored by Crysmann on the grounds that EP clitics also exhbit partly morphological and partly syntactic behaviour: postverbal clitics form a morphological unit with the verb (because they cannot be seprarated from it), while proclitics enjoy more syntactic freedom in that they can be separated from the verb by syntactic material. Crysmann's main claim therefore is that EP clitics must be allowed to be linearised in phrase structure, adopting Kathol's (2000) linearisation-based treatment of German word order in the verbal cluster. Making use of Kathol's model of German 'field theory', it is argued that EP clitic placement can be captured if clitics are ordered into topological fields (cf. Chapter 8). While the enclitic would be positioned within the same 'verbal field', proclitics would have their own topological position.

One further theoretical assumption of Crysmann's analysis is that clitics introduce their own word order domain. Within HPSG, word order domains are generally introduced by words and enable words to be linearised in phrase structure. By assigning this property to clitics, Crysmann

effectively assigns them the freedom to appear in topological fields which may no necessarily be adjacent to the verb.

Upon closer inspection, Crysmann's analysis captures both the preverbal and postverbal position of clitics and it also derives the asymmetry between enclitics and postverbal. Recall that preverbal clitics can be separated from the verb because they occupy different topological fields. It may therefore seem that there is empirical motivation for assigning syntactic 'freedom' to clitics. There are however several problems with the claim that clitics in EP should nejoy such freedom:

a) the analysis is based on analogy between clitics and separable verb-particles, which however is far from obvious, specially because the degree of separability between proclitics and the verb is much more restricted than that of verb-particles. Even though proclitics are in effect separated from the verb, their distribution with respect to the verb is quite local (cf. 19) given that they can only be separated from the verb by a restricted number of non-projecting words (or by a quite reduced amount of two-word units) (cf. Chapter 8). This point seems quite relevant, because it explains why Crysmann must introduce additional constraints to place proclitics within the proximity of the verb. The fact that proclitics turn out to be adjacent to the verb does not seem to fall out from the original assumption that clitics should enjoy syntactic 'freedom'.

In this thesis, we agree with Crysmann in that proclitics are somehow 'liberated' when compared to enclitics. However, we also argue that the proclitic's freedom does not invalidate their purely affixal status. I shall make use of insights by a wide number of authors who have established that 'phrasal affixes' should be regarded as well-established categories. Crysmann's work, on the contrary, assumes that inflectional affixes can only attach to a host in the morphology. It overlooks the numerous studies on phrasal affixation which have convincingly shown that affixes can also attach to a phrasal node (cf. 3.2.2.1). This thesis shows that if a phrasal affix analysis is adopted, both the distribution of proclitics and the scopal effects can be nicely accounted for without introducing new grammatical categories into grammar, such as 'hybrids'. In addition, the local

relation between the proclitic elegantly follows from placement rules which simply state that they are verbal affixes, attaching to a verbal node within VP (cf. Chapter 8).

Crysmann's treatment of clitics as 'hybrids' also poses problems for enclitics. By assigning syntactic freedom to clitics it completely fails to capture the cohering relation between postverbal clitics and the verb. Instead, the unity can only be derived through additional constraints which serve the exclusive purpose joining the word order domain of the clitic with that of the verb, so as to produce one domain (that of the verb-enclitic unit). Unlike in an inflectional analysis, which generates clitics as genuine suffixes and combines them with the verb in the morphology (cf. Chapters 6 and 7), Crysmann's analysis cannot naturally capture the cohering relation between the verb and the enclitic.

Summing up, both Vigário (1999b) and Crysmann (1997, 2000) attempt to accommodate the morphological and syntactic properties of clitic pronouns in different ways: Vigário (1999b) regards clitics as Xº heads (in line with Duarte&Matos 1995) and treats them prosodically like function words with phrasal allomorphs; Crysmann, on the other hand, captures the partly syntactic and partly morphological properties by generatig clitics as morphosyntactic hybrids and linearising them in phrase-structure syntax across topological fields.

As alluded to before, this thesis shares with Vigário and Crysmann the insight that cliticisation in EP cannot be derived though purely syntactic means. It also shares the view that the behaviour of enclitics and proclitics is unequal (cf. also Vigario 1999b, and Duarte&Matos 2000). However, unlike Vigário and Crysmann, I assign genuine suffixal status to enclitics and derive the verb-enclitic unit as a genuine inflected verb form, adopting insights from Miller&Sag (1997) for French, Monachesi (1999) for Italian and Spencer (2000) for Macedonian. Proclitics, on the contrary, are positioned with respect to a phrasal host.

3.3. Summary

This chapter aimed at clarifying the use that shall be made in this thesis of the term 'clitic'. It further aimed at showing that there are different morphological views of cliticisation, one which restricts the morphological properties to clitic-clitic combinations and derives clitic placement as in the syntax (Gerlach 2001a); the other, generally known as the inflectional approach, aims at analysing clitics as verbal affixes, thus attaching clitics to the verb in the morphology. Having compared the claims made by proponents from either type of the approach, it has been argued that the inflectional view offers both a more insightful account of clitic behaviour and a more parsimonious view of grammar. By assimilating clitics to affixes, we make use of already existent categories. In this sense, inflectional studies of cliticisation dispense with the concept clitic altogether, while non-inflectional approach assign theoretical status to clitics and treat as a conceptual category in its own right. Inflectional studies assimilate clitics to affixes and use the term 'clitic' as a pre-theoretic label. In addition, the approach naturally accounts for a wide range of affixal properties displayed by clitics.

This chapter also sketched the EP data and contextualised it within the wider context of Romance linguistics. Throughout the thesis, the comparison between EP and other Romance languages shall be necessary for two reasons: to point the similarities EP shares with other languages (e.g., with respect to enclisis and cluster formation); and also to point out those aspects which motivate an asymmetric analysis of placement.

Part II - Pronominal clitics and verbal inflection

PART II investigates the morphological behaviour of verb-final and verb-internal clitics in EP, and offers a preliminary inflectional analysis of cliticisation.

CHAPTER 4 provides empirical evidence supporting the claim that EP postverbal clitics behave like verbal suffixes. Based on morphological criteria previously used for clitic pronouns in French, Italian and Spanish, it is argued that EP has a fundamentally inflectional clitic system.

CHAPTER 5 concentrates on the phonological behaviour of postverbal clitics and examines previous phonological approaches to cliticisation. Various morphophonological and prosodic phenomena are examined which show that enclitics do not behave phonologically like function words.

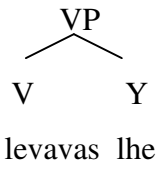
CHAPTER 6 explores an inflectional treatment of suffixing clitics in verb-final and verb-internal position. The intuition underlying this analysis is that clitics are realised as verbal suffixes and combined with the verbal base through inflectional rules. Under this view, the verb-clitic unit is derived as a morphologically cohering unit.

CHAPTER 7 proceeds with the inflectional analysis of suffixing clitics by focusing on the morphophonological effects found both at the verb-enclitic boundary and within the clitic sequence. It is argued that an inflectional treatment of allomorphy provides a more insightful account of the grammatically conditioned nature of non-productive shape alternations.

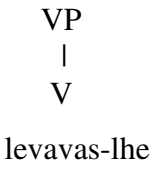
Chapter 4 The affix-status of clitic pronouns

An important question in this study is whether a form like *levavas-lhe* (2sg.asked-3sg.Dat 'you asked him/her') should be regarded as a sequence of two syntactic elements or as one word. In this part of this thesis empirical evidence shall be provided to support the view that the verb-enclitic unit is best represented as in (1b).

(1) a.



b.



Explicit analogies between clitics and affixes shall be made and it is argued, based on morphological and morphophonological properties, that a) enclitics are in morphological construction with the verb and that b) clitic clusters constitute affixal sequences. The chapter is organised as follows: section 4.1 offers a short introduction, section 4.2 focuses on verb-final clitics, while section 4.3 considers verb-internal clitics. Finally, section 4.4 addresses the affixal properties of clitic combinations.

4.1 Introduction

The criteria that shall be used for the identification of the affixal properties of EP clitics are based on a set of well-known diagnostics proposed, in largely theory-independent terms, by Carstairs (1987), Klavans (1980, 1985), Sadock (1991) and Spencer (1992) and, most influentially by Zwicky&Pullum (1983) and Zwicky (1985, 1987). The relevant criteria are given below:

(2) Typical properties of affixes include:

a. High degree of selection with respect to the host

b. Rigid ordering

c. Unpredictable gaps in the combinatorial possibilities

d. Obligatorily repetition on each conjunct

e. Morphophonological idiosyncrasies

These diagnostics, known to some as the 'Zwicky&Pullum criteria', are based on the widely held assumption that morphological rules are characterised by a high degree of idiosyncrasy whereas syntactic rules are defined by their high regularity and predictability. Recent lexicalist-based studies, by Miller (1992), Miller&Sag (1997), Monachesi (1999), and Brines (2001) have applied these diagnostics to languages like French, Italian, and Spanish and shown that clitic pronouns display a wide range of properties that cannot be productively derived in the syntax or phonology. In line with these studies, I apply these tests to EP clitics and investigate the similarities they share with their Romance counterparts.

Anticipating the conclusion, the chapter shows that enclitics i) only attach to verbs, ii) exhibit non-productive allomorphic variation in combination with the verb, iii) cannot be coordinated, iii) undergo allomorphic variation, and iv) trigger idiosyncratic stem allomorphy on the verb (4.2); v) that clitics may intervene between the verb stem and the future/conditional agreement marker (4.3); and that cluster-internal clitics exhibit vi) internal allomorphy, vii) rigid linear ordering, viii) co-occurrence restrictions, ix) syncretism and x) portmanteau morphology (4.4).

4.2 The verb-clitic combination

4.2.1 High degree of selection

Affixes typically attach to items belonging to a specific class of words (Carstairs 1987; Sadock 1991, Spencer 1991, Zwicky&Pullum 1984). Therefore they are generally classified on the basis of

the category of the base they attach to: if they attach to nouns they are classified as nominal affixes, if they combine with verbs, e,g. like the 1st plural marker *–mos* in EP, given in (4a), they are defined as a verbal affixes. Along these lines, the association of an agreement suffix to verbal stems rules out their combination with an adjectival base as in (4b).

(3) a. visita -mos, bebe -mos, fingi -mos
visit -1pl, drink -1pl, pretend -1pl
'(we) visit, (we) drink, (we) pretend'
b. *feliz -mos
happy -1pl
'(we) happy'

The property of selectivity is also found with EP clitics. Like verbal suffixes, enclitics can only attach to verbs:

(4) a. *A criança deu o livro -lhes.
the child gave the book -dat.3pl
'the child gave them the book'
b. A criança deu-lhes o livro.

In fact, enclitics only surface if there is a verb in the clause. If a verbal host is not available (cf. 5b), then object pronouns simply don't surface and a strong pronominal form is used instead (cf. 5c).

(5) a. O trabalho parece -me satisfatório.
the work seems -1sg.dat satisfactory
'the job seems satisfying to me'
b. *O trabalho satisfatório -me.
the work satisfactory -1sg.dat
'the job satisfies me'

c. Um trabalho satisfatório para mim.
a job satisfying for me
'a satisfying job for me'

Object pronouns in the other Romance languages behave in identical ways. They must be adjacent to the verb, and cannot follow or precede any other words belonging to different categories (Monachesi 1999).

On the contrary, word units generally disregard the category of the adjacent words they occur with because the syntactic category of their host is irrelevant, as long as the structural criteria are met. In EP, the preposition *para* 'for' can occur next to a variety of word categories, provided they appear in NP initial position:

(6) a. para crianças 'to or for children' (*para*+noun)
b. para três bibliotecas 'to or for three libraries' (*para*+numeral)
c. para as escolas 'to or for schools' (*para* +determiner)
d. para já 'for now' (*para* +adverb)
e. para mim 'to or for me' (*para* +pronoun)

As a result of the close relation between enclitics and the verb, enclitics are severely restricted in their distribution. As mentioned earlier, they cannot constitute an utterance on their own, they cannot be topicalised, coordinated or modified (cf. ch.3). Instead, enclitic pronouns are in strict adjacency with the verb and nothing can intervene between them. In (7), the adverb *ontem* 'yesterday' can occur before the verb and after the pronoun, but it cannot break up the verb-enclitic unit, indicating that the verb-enclitic unit behaves effectively like a morphological word.:

(7) a. [Ontem] vi -o
yesterday saw -acc.3sg.masc
b. Vi -o [ontem]
saw -acc.3sg.masc yesterday

c. *Vi [ontem] -o
saw yesterday -acc.3sg.masc
'Yesterday, I saw him'

In this respect, clitic pronouns and affixes are very similar, given that affxes cannot be separated from their base by intervening elements: the 1st plural marker *-mos* in (8) forms with the verb a morphological unit which cannot be broken up by intervening words or pauses.

(8) a. Ontem chegá -mos atrasados ao cinema.
yesterday arrived -1pl late to-the cinema
b. * Chegá -[ontem] -mos atrasados ao cinema.
arrived -yesterday -1pl late to-the cinema
'Yesterday, we arrived late at the cinema'

The inseparability illustrated in (7) can only be accounted for if clitics are analysed as morphological elements, behaving very much like the affix in (8). Under this view, the cohering relation between the verb and the enclitic is the result of the integrity of the verb-enclitic unit. The evidence then supports the view that object pronouns should not be classified as syntactic items.

4.2.2 Coordination

It is widely assumed that ordinary affixes have narrow scope and must therefore be repeated under coordination (Zwicky 1985). The obligatory repetition of affixes is illustrated in (9) where the 1st plural agreement marker *-mos* must appear on each conjoined verb.

(9) a. nós come -mos e bebe -mos
we eat -2pl and drink -2pl
'we eat and drink'
b. *nós come e bebe -mos
we eat and drink -2pl

Syntactic constituents, on the other hand, can be shared straightforwardly across conjuncts. In (10), the noun *casas* 'houses' is the complement of each one of the conjoined verb:

(10) Gosto de comprar e vender casas.
like to buy and sell houses
'I like to sell and buy houses'

The behaviour of clitics in coordinated structures is entirely identical to affixes, as shown in (11) where the enclitic must appear on each of the coordinated verbs :

(11) a. *O João viu -o e cumprimentou na rua.
the J. saw -acc.3sg.masc and greeted in-the street
'John saw him and greeted him in the street'

b. O João viu -o e cumprimentou -o na rua.
the J. saw -acc.3sg.masc and greeted -acc.3sg.masc in-the street
'John saw him and greeted him in the street'

c. *O João viu e cumprimentou -o na rua.
the J. saw and greeted -acc.3sg.masc in-the street
'John saw and greeted him in the street'

Repetition is also mandatory with conjoined analytic tenses:

(12) a. A minha mãe tem -me incentivado e tem -me ajudado.
the my mother has -2sg.dat encouraged and has -2.sg.dat helped
'My mother has encouraged me and has helped me'

b. *A minha mãe tem -me incentivado e tem ajudado imenso.
the my mother has -2.sg.dat encouraged and has helped a-lot
'My mother has encouraged me and has helped a lot'

Enclitics cannot only be shared when two coordinated participles share the same auxiliary. In this context, the enclitic appears on the shared auxiliary and appears only once:

(13) A minha mãe tem -me incentivado e ajudado.
the my mother has -2.sg.dat encouraged and helped
'My mother has encouraged me and helped me a lot'

The inability of enclitic pronouns to have wide scope over conjoined verbs (in both analytic and synthethic tenses) is also pointed out by Monachesi (1999), for Italian; by Brines (2001), for Spanish, and by Miller (1992), for French.

An exception to this general behaviour is found when semantically related verbs are coordinated. In this case, it is possible for proclitics (but not enclitics) in Spanish, Italian and French to have wide scope.

(14) Paul les lit et relit sans cesse. (Fr. Kayne 1975)
Paul them reads and re-reads without stop
'Paul reads and rereads them incessantly'

The fact that proclitics can be shared over conjoined verb suggests to some that 'lit' and 'relit' constitute a compound-like verbal unit. Under this account, (14) should not be analysed as a case of wide scope. The behaviour of clitics in coordinated structures shall be addressed in more detail in Chapter 8 [19]. For now, however, the generalisation that enclitics in EP cannot have wide scope over both conjuncts is correct, and, as pointed out by Miller (1992:157), if clitics must be repeated under coordination, then they must have affix status.

[19] In EP, coordination with proclitics is not just restricted to semantically related verbs, as shall be discussed in Part III.

4.2.3 Paradigmatic gaps

Morphological rules often yield paradigms containing irregularities and arbitrary gaps. This generally happens when one particular combination fails to occur for no principled reason. For example, in English the verb *stride* has no participle form (Matthews 1991). Similarly, in EP impersonal verbs like *haver* 'exist' only have 3sg forms (cf. 15); also, defective verbs such as *banir* 'banish' do not have the 1st person singular form of the Present Indicative nor any of the Subjunctive forms (cf. 16).

(15) paradigm of *haver* 'exist'

	Present	Perfect	Imperfect	Subjunctive	Future	Conditional
1sg	-----	-----	-----	-----	-----	-----
2sg	-----	-----	-----	-----	-----	-----
3sg	Há	houve	havia	houvesse	haverá	haveria
1pl	-----	-----	-----	-----	-----	-----
1pl	-----	-----	-----	-----	-----	-----
3pl	-----	-----	-----	-----	-----	-----

(16) paradigm of *banir* 'banish'

	Present	Subjunctive
1sg	-----	-----
2sg	banes	-----
3sg	bane	-----
1pl	banimos	-----
1pl	banis	-----
3pl	banem	-----

Also in Italian, verbs like *spare* and *solere* don't have a present participle form, and verbs like *splendere* and *prudere* lack the past participle form (Monachesi 1999). Patterns created by syntactic rules, on the other hand, do not usually contain such grammatically conditioned irregularities.

Clitic systems display arbitrary gaps in so far as they do not combine productively with all verbal tenses[20]. Although EP clitics can appear in combination with both finite and non-finite forms (17), arbitrary factors seem to restrict the combination. As (18a) shows, past participle forms cannot constitute clitic hosts (Duarte et al.1995) and that nominal phrases must be used instead (18b). This behaviour is quite unlike that of NP complements which can follow past participle verbs without causing ungrammaticality:

(17) a. Fazer -lhe uma surpresa, seria boa ideia.
make -dat.3sg a surprise would-be good idea
'Giving him/her a surprise, would be a good idea.'

b. Os gatos faziam -lhe compania.
The cats made -dat.3sg company
'The cats kept him/her company.'

(18) a. *Compradas -as, regressámos a casa.
bought -acc.3pl.fem, returned to home
'Having bought them, we returned home.'

b. Compradas as flores, regressámos a casa.
bought the flowers, returned to home
'Having bought the flowers, we returned home.'

Italian also provides good examples of paradigmatic gaps. In this language, not all clitics can combine with present participle verbs: so, whereas dative pronouns can follow a present participle verb, the combination with accusatives is restricted to 1st and 2nd accusatives. Completely ruled out are 3rd plural accusatives, with 3rd singular accusatives appearing only marginally (Monachesi 1999).

So, the paradigmatic gaps found across Romance, including EP, are generally dependent on tense features of the verb or on case/person properties of the pronouns. There is no obvious syntactic

[20] Arbitrary gaps are also found inside clitic combinations where co-occurrence restrictions among

or semantic reason to explain why clitics combine with non-finite forms such as the infinitive, but fail to combine with past participles in EP (18), or why present participles in Italian combine only with some accusative clitics but not with others. Under the assumption that syntactic rules are maximally general, these co-occurrence restrictions are best regarded as being morphologically triggered.

4.2.4 Morphophonological effects

Let us now address the incidence of allomorphy at the boundary between verbs and enclitics. Irregular phonological alternations (i.e. allomorphy) can be classified into three groups (Matthews 1991). Lexically conditioned allomorphy is restricted to an arbitrary group of lexical categories (e.g., *-en* participle formation in English). Grammatically conditioned allomorphy is dependent on properties like conjugation class, number, person (e.g., in EP, the Imperfect marker is *-va* in the first conjugation, but *-i* in the second and third) ; and, finally, phonologically conditioned allomorphy is determined by the phonological properties of specific lexical categories (e.g., English plural formation is determined by the preceding phonological segment on the noun). The verb-enclitic unit in EP exhibits all of these types of morphophonological change, thus strongly supporting the morphological attachment of the enclitic to the verb.

In the case of EP pronouns, we find allomorphic alternations conditioned by lexical, grammatical and phonological principles. Of these three types, the two latter are more predominant: pronominal variation found on 3rd accusatives is either grammatically (cf. 19a) or phonologically induced (cf. 19b); verb stem variation is either determined by lexical (cf. 20a), phonological (20b) or grammatical factors (20c). In addition, the changes take place at the boundary between the verb and

---

clitics are quite common. This aspect of the data is discussed in section 4.4

the enclitic, sometimes affecting both the clitic and the verb (as a kind of reciprocal allomorphy). Anticipating the following discussion somewhat, the data can be summarised as follows:

(19) a. dão **-o** → dão-**no** [dɑ̃w̃-u] → [dɑ̃w̃-nu]
give -3sg.acc.masc
'They give it'

b. faz **-as** → fa-**las**
does -3pl.acc.fem
'S/he does them'

(20) a. quer -o → quer**e**-o
wants -3sg.acc.masc
'S/he want it'

b. damo**s** -vos → damo∅-vos
give -2pl.acc
'We give you'

c. fazemo**s** -los → fazemo∅-los
make/do -3pl.acc.masc
'We make/do them'

An important aspect of the data is that these effects are all obligatory, not optional (as seems to be the case in other varieties of Romance, such as Rumanian (Gerlach 2001a). In addition, they can only be found at the verb-enclitic juncture (unlike in Italian, where some effects appear also across word-boundaries[21]). The exclusive nature of the EP data then indicate that the phenomenon is clitic-specific.

[21] In Italian and Romanian, clitic variation appears not to be always obligatory and may also be found with free determiners (cf. Gerlach 2001a). As this section will show, in EP that generalisation does not hold.

4.2.4.1 Clitic allomorphy

In postverbal position, 3rd accusative clitics change their phonological form. The complete inventory of 3rd accusative allomorphs is given below:

(21)

	accusative clitic pronouns			
	singular		plural	
	Masculine	feminine	masculine	feminine
n-form	-no	-na	-nos	-nas
l-form	-lo	-la	-los	-las
j-form	-/j/o	-/j/a	-/j/os	-/j/as
default form	o	a	os	as

The variation between these four forms is determined by either phonological or grammatical factors, as will be discussed below[22].

The *n*-form of the accusative

The *n*-form of 3rd accusative pronouns is selected either after 3rd plural verb forms (in central and southern varieties) or after any nasal final verb (in more northern varieties). This generalisation applies to both lexical verbs (22) and auxiliaries (23) and is not sensitive to the tense value of the verb. The occurrence of *n*-allomorphs with 3pl forms is exemplified with Present and Imperfect tenses below:

(22) a. *lavam-o → lavam-no [laɑ̃w̃-nu]

wash-acc.3sg.masc

'they wash him'

[22] The *j*-form has hardly been mentioned in the literature, but like the other accusative forms it is morphophonologically conditioned. It appears after verbs ending in –z and –s as in *tu comes-o* → *tu comes-/ju/* 'you.sg eat it'.

b. *lavavam-as → lavavam-nas [lɑvavɑ̃w̃-nɑs̆]

washed-acc.3pl.fem

'they washed them'

(23) a. *têm-o dito → têm-no [tɑ̃j̃ɑ̃j̃-nu] dito

have-acc.3sg.masc said

'they have said it'

b. *tinham-as visto → tinham-nas [tiñɑ̃w̃-nɑs̆] visto

had-acc.3pl.fem seen

'they had seen them'

Diachronically the *n*-form is phonologically triggered by the adjacent nasal diphthong found on 3rd plural verb forms in EP. In (24), the endings *-am* and *-ão* corresponds to the nasal diphthong [-ɑ̃w̃], and *-em* to [-ɑ̃j̃]):

(24) 3pl forms of *levar* 'take', *beber* 'drink, and *mentir* 'lie'

	Present	Perfect	Imperfect	Subjunctive	Future	Conditional
levar	lev**am**	Levar**am**	levav**am**	levass**em**	levar**ão**	levari**am**
beber	beb**em**	beber**am**	bebi**am**	bebess**em**	beber**ão**	beberi**am**
mentir	ment**em**	mentir**am**	menti**am**	mentiss**em**	mentir**ão**	mentiri**am**

The role played by the phonology is even more noticeable in the more conservative varieties of the north where any nasal-final verb form (i.e., 3sg Present Indicative or 2sg Imperative forms) triggers *n*-accusative enclitics[23]:

(25) a. O João tem -no [tɑ̃j̃-nu]

the J. has -acc.3sg.masc

'John has it'

[23] There is only a handful of nasal-final verb forms which are not 3pl, such as *põe* 's/he puts; Put!'; *tem* 's/he has; Have!'; *vem* 's/he comes; Come!'; and *retém* 's/he retains; Retain!'; *propõe* 's/he proposes; Propose!'.

b. põe -no [põ̃j-nu]

put -acc.3sg.masc

'put it '

This regularity however should not suggest that the selection of the nasal allomorph is conditioned by purely phonological factors. As the data in (26) shows, nasal diphthongs on nouns (cf. *João* in 26a) or adverbials (cf. *então* in 26b) do not trigger *n*-selection.

(26) a. ... se o João o viu → *... se o João no *[ʒuɐ̃w̃ nu] viu

... if the J. acc.3sg.masc saw

'... if John saw him'

b. ... se então o vendido → *... se então no [ẽtɐ̃w̃ nu] fizeram

... if then acc.3sg.masc sold

'...if they had then sold it'

There is then no doubt that the distribution of the *n*-form in contemporary Portuguese takes only place at the boundary between verbs and enclitics. It is a category-specific alternations which does not follow from productive rules of phrasal phonology.

The idiosyncrasy of the contexts triggering clitic allomorphy is particularly evident in central and southern varieties of EP where the *n*-form of 3rd accusatives enclitics is in fact restricted to 3pl verb forms, as given in (27-28). In these varieties, nasal-final verbs such as 3sg Present Indicative or 2sg Imperative forms select the default (i.e. vowel-initial) form:

(27) a. Vem -a ver! → *vem-na [vɐ̃ɟ-nɑ]

come -acc.3sg.fem see

'Come to see her!'

b. Põe -os na rua! → *põe-nos [põ̃j-nuš]

put -acc.3pl.masc in-the street

'Throw them out'

In these varieties the alternation is not associated to phonological properties of the preceding word, but determined by specific grammatical features, namely agreement properties of the verb (cf. Vigário 1999b for similar view).

So far, we have seen that the choice of the *n*-accusative form is determined by grammatical features in central/southern varie ties. It can also extend to any nasal final verb forms in the northern varieties where the choice of this particular clitic pronoun is determined instead by verbal phonology. However, note that even in this latter case, it is not clear that the phonology is the only triggering force, for a verb form such as *vim* 'I came' would only very marginally select a nasal allomorph:

(28) Vim -as buscar. → *Vim-nas [vĩ nɑ̃ʃ] buscar.
came -acc.3pl.fem take
'I came to take them'

If this intuition is correct, then the choice between a vowel-initial and an *n*-intial form may not be based purely on an adjacent nasal sound[24]. This issue however is orthogonal to the already established morphophonological status of the allomorph selection.

The *l*-form of the accusative

Another pronominal allomorph of the 3rd accusative paradigm is given in (29). This form is *l*-initial and is selected when the preceding verb ends in one of the following consonants, i.e. -s, -z, -r:

[24] In the remainder of this thesis, I will associate *n*-allomorphs to 3pl verb forms (cf. Vigário 1999b for similar position). However, both uses of the *n*-allomorph (either conditioned by 3pl features or by verb-final nasals) are associated to verb forms only and constitute therefore evidence for morphophonological status.

(29) a. *dizemos-o → dizemo∅-lo
say-acc.3sg.masc
'(we) said it'

b. *faz-o → fá∅-lo
does-acc.3sg.masc
's/he does it'

c. *fazer-as → fazê∅-las
make-acc.3pl.fem
'(to) make them'

Auxiliaries with word-final consonants also select the *l*-form:

(30) a. ter-o dito → tê∅-lo dito
have-acc.3sg.masc said
'(to) have said it'

b. temos-o dito → temo∅-lo dito
have-acc.3sg.masc said
'(we) have said it'

This variation is found between verbs and enclitics, but not across word boundaries. For example, vowel initial words preceded by consonant-final words do not undergo this alternation, as in (31a-b). Likewise, definite articles which are phonologically similar to the 3rd accusative clitics, as in (31c), also block the change.

(31) a. lápis azul → *lápi∅ lazul
'blue pencil'

b. giz usado → *gi∅ luzado
'used chalk'

c. Tu compras o bolo → Tu *compra∅ lo bolo
'You buy the cake'

Similarly, if clitic pronouns precede consonant final words other than verbs, selection of the *l*-form would be ungrammatical:

(32) a. Todos os clientes o viram → *Todos os cliente∅-lo viram
all-the customers acc.3sg.masc saw
'all customers saw him'

b. Eu sei que o giz a irrita → *Eu sei que o gi∅-la irrita
I know that the chalk acc.3sg.fem irritates
'I know that chalk annoys her'

The data then clearly illustrates that the alternation is not a post-syntactic phenomenon[25]. It also shows that the allomorph selection is not determined by grammatical properties of the context, as with the *n*-form. Instead it is triggered by the phonological properties of the preceding verb.

One further aspect about the *l*-form is that it triggers deletion of the verb-final consonant (deletion is signalled with '∅' in 32). This a typical case of stem-allomorphy which shall be addressed in detail in section 4.2.4.2.

[25] In Old Portuguese, this dissimilation is also found between consonant-final function words and vowel-initial determiners, giving rise to lexicalised prepositions, as in (i). In contemporary Portuguese, these effects however are not productive anymore (Barbosa 1996).

(i) por a → pela
'through the.fem'

Glide-initial allomorph

Before I move on to the default realisation of the accusative clitic, I shall briefly refer to the glide-initial form -[j] o(s), -[j] a(s). This is a substandard allomorph, only found in colloquial Portuguese:

(33) a. diz -o → diz-[ju]
says -acc.3sg.masc
's/he says it'

b. lavas -as → lavas - [jɑs]
clean-acc.3pl.fem
'you clean them'

This type of glide is not inserted to break up a hiatus, as happens so frequently in EP, but it simply appears after a verb-final consonant which is thereby incorporated it into the onset of the enclitic syllable. What is also interesting about this glide is that it does not occur elsewhere in the language, neither with function words (cf. 34) nor content words (cf. 35).

(34) a. as uvas *[ɑzjuvɑš] → [ɑzuvɑš]
'the grapes'

b. todas as ... *[todɑzjaš] → [todɑzɑš]
'all the …'

(35) belas asas *[bɛlɑzjazɑš] → [bɛlɑzazɑš]
'nice wings'

This type of allomorphy is clearly conditioned by the phonological make-up of the preceding verb, and replaces the use of the *l*-allomorph in similar contexts.

The default accusative form

The default realisation of the 3rd accusative pronoun takes place when none of the previous contexts meet, that is, when the verbs preceding the clitic pronoun are neither marked for 3rd plural nor do

they end in -s,-z,-r. The verbs which generally occur with vowel initial 3rd accusative clitics are diphthong-final (36) or vowel-final (37a):

(36) a. lavei -a
washed -acc.3sg.fem
'I washed her'

b. levou -as
took -acc.3pl.fem
's/he took them'

In vowel-final verb forms, the hiatus that is formed by the adjacent vowels is broken up by the insertion of glide (cf. 37a). This phonological process is mandatory word-internally (37a-b), but optional elsewhere (cf. 37c):

(37) a. comi-a [ija]
ate.Impf.1sg-acc.3sg.fem
'I ate-it'

b. comia[ija]
ate.Pret.1sg almonds
'I ate'

c. comi amêndoas ([ija],[ia])
ate.Impf.1/3sg
I ate almonds'

To conclude this overview about clitic-allomorphy, one final word is in order about the the 3rd singular form of *querer* ‘want’, in (38):

(38) a. O Paulo quer o livro
‘Paul wants the book'

b. O Paulo *quer-o
‘Paul wants him/it’

c. O Paulo *que-lo

d. O Paulo quer**e**-o

As (38a) shows, the verb form is consonant-final. So we expect that it cannot combine with a vowel-initial enclitic (cf. 38b), and that it must select instead the *l*-form (cf. 38c). This option however is unexpectedly blocked and what we find is an epenthetic vowel which is added to the verb form and triggering the selection of the default accusative clitic in (38d). Elsewhere, however, consonant-final forms of QUERER select the *l*-form, as expected:

(39) querer-o → querê∅-lo
(to) want it

The fact that a vowel is added to the verb stem appears then to indicate that this is instead a case of lexically and grammatically triggered stem-variation.

To sum up the discussion about 3rd accusative clitics, the following table illustrates the co-occurrence of the three standard allomorphs:

(40)

Perfect Preterite *comer* 'eat'	
Non-cliticised paradigm	verb-enclitic forms
eu comi	eu comi-**o**
tu comeste	tu comeste-**o**
ele comeu	ele comeu-**o**
nós comemos	nós comemo-**lo**
vós comestes	vós comeste-**lo**
eles comeram	eles comeram-**no**

In (40), the *n*-form appears with the 3pl verb form, the *l*-form is selected by the three consonant-final forms and the vowel initial form is placed on the remaining verb forms. On the basis of this data, it was shown that the shape variations displayed by the clitic pronoun are dependent on grammatical and phonological properties of the preceding verb form. This type of sensitivity displayed by clitics is an indication of their status as genuine affixes.

4.2.4.2 Stem-variation

As mentioned before, enclitics not only undergo allomorphic alternations but also induce variation on the verb. In this section, then, we shall examine clitic-induced stem-allomorphy.

There are two sets of clitics which trigger morphophonological changes to verbs, namely

- *l*-accusatives, and
- 1st /2nd plural clitic pronouns,

Both contexts induce verb-final consonant deletion on the preceding verb.

The accusative context

As mentioned earlier, the *l*-form of 3rd accusative clitics is triggered by consonant-final verbs. In (41), it is shown that the selection of this particular clitic allomorph has an allomorphic effect on the verb: in particular, the verb-final consonant preceding the *l*-accusative is deleted. This generalisation applies to all consonant-final verb forms, regardless of tense or category (deletion is signalled with '∅'):

(41) a. *levávamos-o → levavamo∅-lo
took-acc.3sg.masc
'(we) took it'

b. *tinhas-o dito → tinha∅-lo dito
had-acc.3sg.masc said
'(you.sg) had said it'

Other accusative clitics fail to trigger this effect:

(42) a. vimos-te → *vimo∅-te
saw-acc.2pl
'we saw you'

b. vistes-me → *viste∅-me
saw-acc.2sg
'you.sg saw me'

Deletion also fails to take place across word boundaries, for example when a consonant-final verb is followed by an *l*-initial word, as in (43). The phenomenon is also blocked word-internally (cf. 45):

(43) a. compramos luvas → a'.* compramo∅ luvas
'we bought gloves'
b. diz logo → b'. *di∅ logo
'speak later'

(44) legislar → *legi∅lar
'legislate'

(43) then indicates that the phonological change suffered by the verb does not follow from phrasal phonology; (44) shows that the variation is restricted to the verb-enclitic unit, being triggered by the phonological properties of one specific 3rd accusative enclitic.

The 1st / 2nd plural context

The second case of consonant-final deletion is triggered by 1st and 2nd clitic pronouns (regardless of case or reflexivity):

(45) a. Nós *vêmos-nos → vêmo∅-nos
we see-1pl.reflexive-acc
'we see ourselves'
b. Nós *davamos-vos → davamo∅-vos
we give-2pl.dative
'we give you'
c. Nós encontramos-vos → encontramo∅-vos
we met.2pl.acc
'we met you'

This deletion only takes place if the preceding verb is 1st pl. Other consonant final verb forms, like the 2nd singular form in (46) or the 2rd sg form in (47) will not trigger the alternation:

(46) Tu recebes-nos → *recebe∅-nos
you.sg receive-acc.1pl

(47) Ele quis-nos → *qui∅-nos
you.pl receive- acc.1pl

This type of stem-allomorphy then is clearly grammatically conditioned: only 1st and 2nd plural clitics trigger deletion and only on 2pl verbs. The variation is therefore dependent on the person and number features of both the enclitic and the verbal host. Any attempt at accounting for the data through productive phonology would completely fail to capture the inherently morphological status of the phenomenon.

Vowel epenthesis

In addition to segment deletion, we also find that the opposite can happen. As referred to above, 3sg Present Indicative forms of QUERER 'want' undergo vowel epenthesis when followed by a 3rd accusative enclitic (48):

(48) quer-(e)-os → *que∅-los
'I want them'

As mentioned earlier, the insertion blocks the selection of the *l*-form of the accusative clitic and determines the selection of the default, vowel-initial clitic. This process occurs only with this particular verb form, since other consonant-final forms of the lexeme behave as expected (49).

(49) querer-os → quere∅-los
'(to) want them'

As mentioned earlier, the motivation for vowel epenthesis appears not to be phonological. The variation suffered by the 3sg form then is clearly a case of lexically and grammatically conditioned alternation. Lexical, on the one hand, because the phenomenon appears to be restricted to one single verb, and grammatical, on the other, because it selects a specific person and number combination. Such a combination of idiosyncratic factors would make it quite difficult to derive the data though productive phonological rules.

Summing up, three properties can be attributed to the shape variations discussed in this section: a) they are obligatory, b) they only apply inside the verb-enclitic unit, and c) they are determined by a combination of grammatical, phonological or lexical features. This highly idiosyncratic nature of clitic and stem allomorphy therefore indicates that enclitics are in morphological construction with the verb.

4.2.5 Summary

The main goal of this section was to argue that enclitics undergo all of the Zwicky&Pullum criteria listed in (1). From the interaction between the enclitic and the verb, it was shown that enclitics behave like verbal suffixes. They are a) adjacent to the verb and inseparable from it (4.2.1), b) they cannot have wide scope over coordination (4.2.2) and c) they are affected by arbitrary co-occurrence restrictions (4.2.3), and d) both the enclitic and the preceding verb undergo phonological shape variations (4.2.4) Having provided evidence to support the view that clitics are in morphological construction with the verb, I will now examine the phenomenon frequently referred to as mesoclisis.

4.3 Mesoclisis

Further evidence for affixal-status is provided by verb-internal clitics. This clitic-position is rarely found in the other Romance languages. In EP it is also restricted to future and conditional verb forms, as illustrated in (50):

(50) mesoclitic placement in the future and conditional paradigms of *levar* 'take'

	future indicative	conditional
1sg	levar-te-ei 'I will take you'	levar-te-ia 'I would take you'
2sg	levar-me-ás 'you will take me'	levar-me-ias 'you would take me'
3sg	levar-me-á 'she will take me'	levar-me-ia 'she would take me'
1pl	levar-te-emos 'we will take you'	levar-te-íamos 'we would take you'
2pl	levar-me-eis 'you will take me'	levar-me-íeis 'you would take me'
3pl	levar-me-ão 'they will take me'	levar-te-iam 'they would take you'

As the table shows, what characterises these cliticised verb forms is the fact that the clitic suffix is positioned between the infinitival stem and the tense/agreement suffixes (Leeuw 1995). The placement of clitics in verb-internal position is traditionally referred to as mesoclisis, and it can also be found in auxiliary-verb constructions:

(51) a. As crianças ter -nos -ão visto.
the children have -1pl.acc -fut.3pl seen
'The children will have seen us'

b. As crianças ter -lhes -iam agradecido.
the children have -1pl.acc -cond.3pl thanked
'The children would have thanked us'

Mesoclisis has been subject to some debate because of its double stress which appears on the theme vowel of the stem and the future/agreement marker. This is quite untypical in EP verb forms (cf. Chapter 5 for details). But mesoclisi has also sparked interest because the future and conditional

marker do not occupy the expected postverbal position as in their Romance counterparts. These facts have suggested to some that the verbs forms in (50) do not constitute one synthetic verb form but should instead be derived as analytic verbal constructions (Mateus 1983, Vigário 1999b).

For example, in Mateus (1983) the representation of word-internal boundaries on mesoclitic verb forms does not seem to treat clitics as verbal affixes. As van der Leeuw (1997) points out, Mateus marks a word with # and an affix with +, while the whole morphological structure is marked as ##:

(52) ## bat # e # r + # te # + á + s ## (Mateus 1983)
'(you) will beat yourself'

Similarly, Vigário (1999a,b) does not regard verb-internal clitics as part of an inflected verb form, but regards them as function words (cf. Chapter 5).

Despite the long-standing arguments against an inflectional analysis of mesoclisis, there is a wide range of evidence seriously weakening a syntactic approach to clitic placement as proposed by Vigário (1999b) and Gerlach (2001a). In what follows, I will briefly show which problems are raised by a non-inflectional view of mesoclisis, by examining the behaviour and properties the tense and agreement markers (4.3.1), the verb-internal clitic (4.3.2) and the infinitival verb form (4.3.2) [26].

4.3.1 Tense and agreement markers

Under a non-inflectional analysis of mesoclisis, the verb forms in (50) are regarded as comprising three syntactic word-level units: the verb, the enclitic and the tense/agreement unit. Starting with the

[26] The phenomenon of 'mesoclisis' is discussed in various parts throughout this thesis. Because of the importance of the phonological aspects of mesoclisis, I address these in a separate chapter (cf. Chapters 5 and 6 for a discussion of phonological properties; cf. also Chapters 6 and 7 for more detailed arguments against previous analyses of mesoclisis).

tense and agreement suffixes, let us see why they should be viewed as ordinary verbal suffixes. (cf. also Chapter 6 for further discussion)

a) homophonous suffixes

First, tense markers and the agreement markers in ordinary (future and conditional) verb forms and cliticised verb forms are exactly identical:

(53) future tense suffixes and agreement markers

	Future indicative	
1sg	levar-**e.i**	levar-lhe-**e.i**
2sg	levar-**á.s**	levar-lhe-**á.s**
3sg	levar-**á**	levar-lhe-**á**
1pl	levar-**e.mos**	levar-lhe-**e.mos**
2pl	levar-**e.is**	levar-lhe-**e.is**
3pl	levar-**a.N**	levar-lhe-**a.N**

(54) conditional suffixes and agreement markers

	Conditional	
1sg	levar-**ia**	levar-lhe-**ia**
2sg	levar-**ia.s**	levar-lhe-**ia.s**
3sg	levar-**ia**	levar-lhe-**ia**
1pl	levar-**ía.mos**	levar-lhe-**ía.mos**
2pl	levar-**íe.is**	levar-lhe-**íe.is**
3pl	levar-**ia.N**	levar-lhe-**ia.N**

The tables show that future and conditional forms have two tense markers, namely *-e* and *-a* for future and *-ia* and *-ie* for conditional (tense and agreement suffixes are highlighted). Note that the same variation is found in both cliticised and non-cliticised verb forms. Likewise, agreement suffixes found in (53) and (54) are also homophonous. So, the identity between cliticised and non-cliticised verbs extends both to the tense marker, which immediately follows the mesoclitic, and to the agreement marker (if present), which follows the tense marker. In addition, both suffix positions occur in the same order regardless of whether the verb combines with a clitic or not. In addition, the

agreement markers (i.e., *-s*, *-mos*, *-N*) in mesoclitic verb forms are also identical to the agreement markers on most of the verb forms of the EP paradigm (Cunha&Cintra 1987)[27].

Under a non-inflectional account, the wide-ranging similarity between both sets of suffixes - i.e., the tense markers and agreement markers – would be completely accidental. Either a syntactic analysis or a 'compounding' analysis would fail to explain why the set of suffixes on cliticised verbs is the same. The observed similarity can only be captured through an inflectional analysis which assumes that tense and agreement suffixes in EP can appear either after the theme vowel or after a clitic suffix (cf. Chapters 5 and 6 for further discussion).

b) distribution of verb-final suffixes

There is also a strong distributional similarity between the tense and agreement suffixes occurring in ordinary future/conditional verb forms and in cliticised verb forms: they don't have wide scope over coordinated verbs (55a), they don't coordinate (55b) and they don't undergo subject-auxiliary inversion like ordinary verbal suffixes (55c). This, of course, is precisely what one would expect of verbal affixes.

(55) a. *lavar -lhe e limpar -lhe -emos a casa
wash -3sg.dat and clean -3sg.dat -fut.2pl the house
'We will wash and clean the house for him/her'

b. *lavar -lhe -ei ou -ia a roupa
wash -3sg.dat -fut.1sg or cond.1sg the clothes
'I will or would wash the clothes for him/her'

c. *ei -lavar -lhe a roupa
fut.1sg - wash -3sg.dat the clothes
'I will wash his/her clothes for him/her'

[27] I follow Pereira (1999) in representing 3rd plural agreement as a nasal sound. In (53-54), -N signals the nasalised dipthong in future and conditional forms.

Under the view that tense/agreement markers in cliticised future/conditional verb forms are syntactic units, this fact turns out to be mere accident. As syntactic units (more precisely auxiliary units, Vigário 1999b) one would expect them to be more mobile or syntactically active like the auxiliaries in (55). However, this is clearly not the case.

(56) a. O João tinha comido e dormido.
the J. had eaten and slept
'J. had eaten and slept'
b. Eles teriam ou terão visto alguma coisa.
They would-have or will-have seen some thing
'They might have seen or (really) have seen something'
c. Teria ele visto alguma coisa?
had he seen some thing
'Might he have seen something?'

c) semantic distinction

A further interesting aspect is based on the semantic differences between ordinary auxiliaries and the tense/agreement endings on mesoclitic verb forms. Raposo (2000: 284) claims that the Future auxiliary can be used independently, as in (57):

(57) Penso que ele há-de cantar
think.1sg that he has-to sing
'I think that he will sing'

Closer inspection however shows that the synthetic Future and *haver de* + Infinitive are distinct constructions. Likewise for the synthetic Conditional and the Imperfect auxiliary havia de + Infinitive, as in *Ele havia de cantar* 'he should sing'. They are distinct stylistically and there are subtle modal differences in the semantics. Not also that the verbal stem in the *haver de* + Infinitive

construction is not always identical to the verbal stem in the synthetic form, e.g. *havia de fazer* 'had to do' vs. *faria* 'would do'.

d) agglutinated suffixes

There is one further argument in favour of viewing the tense and agreement markers in 'mesoclitic verb forms' as genuine suffixes. Note that syntactic and weak-syntactic treatment are forced to regard the tense/agreement markers in cliticised future and conditional verb forms as an auxiliary unit. Even though there is historical motivation for this claim, it fails to explain why the auxiliary unit in the mesoclitic verb form is phonologically and semantically exactly identical to the tense+agreement suffixes of the non-cliticsed future and conditional forms. This similarity extends both to verb forms with discrete tense and agreement suffixes (e.g., 1st plural verb forms) and to verb forms whose tense and agreement markers are realised by one single suffix (e.g., 1st singular and 3rd singular). Under the non-homogenous view, this similarity is treated as a purely accidental phenomenon.

Summary

Summing up, then, even though there is historical motivation for regarding the tense/agreement endings on cliticised future/conditional verb forms, such as *-emos*, as an auxiliary unit, there is insufficient motivation in Modern EP for maintaining this claim. The arguments present above show that there are empirical arguments for regarding tense and agreement suffixes in (53-54) as comprising a sequence of ordinary future and conditional tense and agreement suffixes.

The approach proposed by Duarte et al. (1995) and Duarte&Matos (2000) assumes an intermediate claim according to which the Future/Conditional marker has become a "T-affix", that is, that is has become "lexicalized under T" (a Tense node in the syntax). It is difficult to find a clear interpretation for this claim because of the inexplicit formulation. It may amount to saying that they are neither inflectional nor syntactic. If that is the case then they are regarded as special affixes,

contrary to evidence which shows that they should be regarded as ordinary affixes (cf. discussion in sections 4.3.1 a, b, c). On the other hand, if it is assumed that we are dealing after all with an affixal inflection, then it remains to be explained how non-cliticised future and conditional verb forms are derived. If they are analysed without a lexicalised T-affix, then it means that the analysis assumes two types of tense and agreement markers. Duarte et al (1995) and Duarte&Matos (2000) however leave this topic unaddressed.

4.3.2 The clitic suffix

The inflectional status of the cliticised verb-forms in (50) is further supported by the morphological behaviour of clitics. The pronominal clusters behave exactly like affixes in that they both trigger and undergo idiosyncratic stem allomorphy. These facts are generally left unaddressed in syntactic analyses (e.g. Raposo 2000, Duarte and Matos 2000) or derived through 'special' phrasal phonology in weak-syntactic approaches (Geralch 2001, Vigário 1999b).

The affixal properties that will be summarized below have been already addressed in section 4.2, with respect to verb-final clitic suffixes:

- Strict adjacency to the verb (4.2.1):

(58) dizer -lhe -emos amanhã (not: *dizer-[amanhã]-lhe-emos)
say -dat.3.sg -Fut.1pl [tomorrow]
'(we) will tell him/her tomorrow'

- Narrow scope over coordination (4.2.2):

(59) a. cumprimentá -los -emos e beijá -los -emos
greet -acc.3pl.masc -Fut.1pl and kiss -acc.3pl.masc -Fut.1pl
'(we) will greet them and kiss them'
b. *cumprimentá -los -emos e beijaremos
greet -acc.3pl.masc -Fut.1pl and kiss.Fut.1pl

(60) a. * cumprimentaremos e beijá -los -emos
greet.Fut.1pl and kiss -acc.3pl.masc - Fut.1pl

- Morphophonological variations (e.g. undergo and induce shape variation) (4.2.4):

(61) lavar -o -emos → lavá∅-lo-emos
wash- -acc.3sg.masc -fut.1pl
'we will wash him'

- Portmanteau clusters (cf. 4.2.4)

(62) lavar -lho -emos
wash -dat.3sg/acc.3pl.masc -Fut.1pl
'we will wash it for him/her'

The properties, then, show that there is sufficient empirical motivation for regarding mesoclitics as verbal suffixes. In this thesis, mesoclitics and enclitics will be regarded as one and the same affixal unit.

4.3.3 The verbal base

Having addressed the behaviour of both tense/agreement markers and clitic suffixes, I will briefly consider the status of the verbal base to which mesoclitics and tense/agreement suffixes attach. This point will be addressed in more detail in Chapter 6. For now, I will only point out that the infinitival verbal base in (50), which constitutes the host for verb-internal clitics, is exactly identical to the verbal base to which ordinary future/conditional tense suffixes attach in non-cliticised verb forms (compare the ordinary future and conditional verbs with the cliticised counterparts in 53 and 54). The base is given below in italics:

(63) a. *levar*-e-mos
b. *levar*-lhe-e.mos

This observation is rather trivial but has important implications for an analysis of mesoclisis.

Any approach deriving mesoclisis through non-inflectional means is forced to assume that there are two different 'infinitival bases' in (53 and (54), one for cliticised verbs and another for non-cliticised. The reason why two bases would be needed is simple. In synthetic future verb forms, it is well-known that even though the infinitival stem is homophonous with the infinitive verb, it does not carry any [-finite] features (in fact, it would make little sense to combine a non-finite verb with finite tense suffixes). Stem that are formally (but not semantically) identical to fully inflected verb forms are called 'parasitic stems' or 'Priscianic stems' (Matthews 1972, Aronoff 1994). Future and conditional stems in EP are of this type. However, this idea can only be captured if the future and conditional tense/agreement affixes are combined with the stem in inflectional morphology, because that is the component in which stems are available. Now, if cliticised future verb forms are combined in the syntax, the verbal base must be regarded as a word form (presumably an infinitive verb) serving as the host for clitics and the auxiliary unit. However if the verbal base is not regarded as an infinitival stem, it is far from clear how the 'future' or 'conditional' meaning of the verb forms will be derived; in addition, there is no evidence for a non-uniform treatment of the verbal base in synthetic verbs and in cliticised verb forms.

4.3.4 Summary

Various arguments were given to support the view that the structures in (50) constitute morphologically cohering verb forms. As to tense suffixes, they are formally and distributionally exactly identical to the endings on non-cliticised future and conditional verbs; likewise the regular agreement endings are the same throughout most of the EP verbal conjugation. As to clitics, they display all the affixal properties found with clitics in verb-final position. Finally, the infinitival verb form constitutes an infinitival *stem* , thus indicating that the combination between stem-clitic and

tense/agreement can only be derived as an inflectional process. The question then is how to capture the fact that clitic suffixes appear verb-internally. This is an issue that will concern us in detail in Chapter 6, where 'mesoclisis' is treated as an instance of suffix-reordering.

4.4 The clitic cluster

This section considers the morphological behaviour of clitics inside the clitic string. Clitic pronouns in EP (and in Romance, in general) combine into clusters. In EP, clusters comrise at most two clitics (see below):

(64) A criança deu -mo. (=me+o)
the child gave -dat.1sg/acc.3sg.masc
'the child gave it to me'

In this section I examine the following cluster-internal regularities, arguing that they further support the affixal status of clitics, posing problems to syntactic analyses.

a) rigid ordering of clitics inside the cluster
b) co-occurrence restrictions between clitics inside the cluster
c) clitic induced morphophonological effects (i.e. allomorphic alternations, fusional morphology and syncretism) at the clitic-clitic boundary

Similar idiosyncrasies have been reported for Italian (Monachesi 1999), for French (Miller&Sag 1997), for Spanish (Brines 2001), Cf. also Crysmann (2002), for EP.

4.4.1 Rigid clitic order

Let us first recall the complete pronominal clitic paradigm of EP. As mentioned earlier, there are reflexive, dative and accusative clitics:

(65)

	1 sg	2 sg	4 sg		1 pl	2 pl	4 pl	
			fem	masc			fem	Masc
accusative	me	te	a/-la/-na	o/-lo/-no	nos	vos	os/-los/-nos	as/-las/-nas
dative			lhe				lhes	
reflexive			Se				se	

The first observation about multiple occurrences of the clitics in (65) is that not all logical combinations are possible. The order in which clitics can appear is seriously restricted, as the template in (66) shows: there is the Reflexive-Dative cluster where reflexives must precede datives (66a), and the Dative-Accusative cluster where datives must precede accusatives (66b).

(66) a. REFL>DAT
b. DAT>ACC

Such ordering constraints are quite unlike the rather free distribution of the corresponding nominal phrases. If we take, for example, the combination between direct objects (i.e. accusative) and indirect objects (i.e. dative) in (67), we see that the dative NPs can either precede the accusative NP (67b) or follow it (67a). With clitics, on the contrary, accusatives must always follow datives (cf. 66b), as illustrated in (67c).

(67) a. A criança deu o livro ao vizinho. (Subj > Acc > Dat)
the child gave the book to-the neighbour
'the child gave the book to the neighbour'

b. A criança deu ao vizinho o livro. (Subj > Dat > Acc)
the child gave to-the neighbour the book
'the child gave the neighbour the book '

c. *O João deu -as -vos.
the J. gave -acc.3pl.fem -dat.2pl
'João gave them to you'

The distributional difference between nominal phrases and clitics suggests that clitic ordering does not follow from general principles of syntax. The same generalisation has been made about clitic order in other Romance languages (Perlmutter 1970, Monachesi 1999, Gerlach 2001a). Rigid ordering then is a property which fails to support the syntactic status of clitics. The introduction of syntactic processes to account for a narrow selection of categories is not what one would expect of a maximally general syntax. On the contrary, affixes typically combine with the base in a fixed linear order (Stump 1993, 2001; on Swahili and Bulgarian).

4.4.2 Co-occurrence restrictions

In addition to the above linearisation constraints, clitic clusters are also subject to feature co-occurrence restrictions. As shown in (68-69), within the reflexive-dative cluster, only 3rd reflexives can combine with dative clitics.

(68) A professora apresentou -se -me.
the teacher introduced -refl.3sg -dat.1sg
'the teacher introduced herself to me'

(69) *Eu apresentei -me -lhe.
I introduced -refl.1sg -dat.3sg
'I introduced myself to him/her'

Again, there appears to be no syntactic or semantic motivation for this restriction, since reflexive clitic can combine freely with a direct object phrases. In (70), a 2nd person reflexive co-occurs with a direct object complement:

(70) Apresentei -me a ele.
(I) introduced -refl.1sg to him
'I introduced myself to him'

As to the dative-accusative cluster, we observe that dative pronouns can only combine with 3rd person accusative clitics:

(71) a. Mandei -vo -lo. (not : *vos+o)
(I) sent -dat.2pl -acc.3sg
'I sent him to you'

b. *Entregámos -lhe -te.
(we) delivered -dat.3sg -acc.2pl
'we handed you over to him'

Again, no such restriction applies with full pronominal phrases:

(72) Entregámos -te a ele
(we) delivered -acc.2sg to him
'we handed you over to him'

Common to the reflexive-dative cluster and also to the dative-accusative cluster is the constraint on first and second person pronouns which can never co-occur, regardless of case:

(73) a. *O João apresenta -me -te (*1/2)
the J. introduces -dat.1sg -acc.2sg
'John introduces me to you'

b. *O João apresenta -mo
the J. introduces -dat.1sg/acc.3sg-masc
'John introduces him to me'

Again, no such restriction applies with full pronominal phrases:

(74) O João apresenta -te a mim.
the John introduces -2sg.dat to me
'John introduces yourself to me'

A similar restriction on the co-occurrence of 1st person clitics with 2nd person clitics is found in the other Romance languages (Monachesi 1999, Gerlach 2001a).

As a result of the combinatorial constraints and ordering restrictions applied to clitic combinations, the complete set of clitic clusters in EP is as follows:

(75) Clitic clusters of type 'Ref-Dat'

	1sg.dat	2sg.dat	3sg.dat	1pl.dat	2pl.dat	3pl.dat
3.sg.refl	se-me	se-te	se-lhe	se-nos	se-vos	se-lhes
3.pl.refl	se-me	se-te	se-lhe	se-nos	se-vos	se-lhes

(76) Clitic clusters of type 'Dat-Acc'

	1sg	2sg	3.sg.m.acc	3.sg.f.acc
1sg.dat	---	---	mo (me+o)	ma (me+a)
2sg.dat	---	---	to (te+o)	Ta (te+a)
3sg.dat	---	---	lho (lhe+o)	lha (lhe+a)
1pl.dat	---	---	no-lo (nos+o)	no-la (nos+a)
2pl.dat	---	---	vo-lo (vos+o)	vo-la (vos+a)
3pl.dat	---	---	lho (lhes+o)	lha (lhes+a)

	1pl	2pl	3pl.m.acc	3sg.f.acc
1sg.dat	---	---	mos (me+os)	mas (me+as)
2sg.dat	---	---	tos (te+os)	tas (te+as)
3sg.dat	---	---	lhos (lhe+os)	lhas (lhe+as)
1pl.dat	---	---	no-los (nos+o)	no-las (nos+as)
2pl.dat	---	---	vo-los (vos+o)	vo-las (vos+as)
3pl.dat	---	---	lhos (lhes+os)	lhas (lhes+as)

The tables in (75) and (76) clearly show that many potential combinations are not available. The motivation underlying these gaps, as mentioned earlier, cannot be related to syntactic function or semantic role since the feature combinations which are excluded in clitic clusters can be expressed in syntax by full pronominal forms. Idiosyncratic combinatorial restrictions however are typical of inflectional systems, as illustrated in Stump (2001, Chapter 2).

4.4.3 Morphophonological effects

Morphophonological effects are quite productive inside the clitic cluster (cf. 4.2.4, for allomorphic alternations at the verb-enclitic boundary). As will be shown next, some of the regularities found cluster-internally include

a) pronominal allomorphy (4.4.3.1)
b) portmanteau formation (4.4.3.2), and
c) syncretism (4.4.3.3).

4.4.3.1 Cluster-internal allomorphy

In the DATIVE-ACCUSATIVE cluster, 1st and 2nd plural dative clitics trigger pronominal allomorphy on 3rd accusative clitics and undergo themselves consonant-deletion:

(77) a. A Maria *compra -nos -o.
b. A Maria compra -no∅ -lo.
the Maria buy -dat.3sg.masc -acc.3sg.masc
'Maria buys it for us'

(78) a. As crianças *dão -vos -as.
b. As crianças dão -vo∅ -las.
The children give -dat.2pl -acc.3pl.fem
'The children give them to you'

In (77) the clitic clusters comprise a 1st plural dative clitic followed by a 3rd singular masculine accusative, and in (78), the dative clitic is 2nd plural and the accusative clitic is 3rd plural feminine. In common, both examples illustrate that the dative clitics trigger selection of the *l*-allomorph and, at the same time, the *l*-allomorph triggers word-final consonant deletion on the preceding dative clitic. This alternation happens with any of the 1st and 2nd plural person dative+accusative combinations, as the inventory in (79) shows:

(79)

	3sg.m.acc	3sg.f.acc	3pl.m.acc	3pl.f.acc
1pl.dat	(nos+o)→ no-lo	(nos+a)→ no-la	(nos+os)→ no-los	(nos+as)→ no-las
2pl.dat	(vos+o)→ vo-lo	(vos+a)→ vo-la	(vos+os)→ vo-los	(vos+as)→ vo-las

The shape variation suffered by the accusative clitic is reminiscent of the morphophonological variation found between consonant final verbs and 3rd accusative clitics (cf. 4.2.4 above):

(80) *compramos -o → compramo∅-lo
(we) bought -acc.3sg.masc
'we bought it'

What this shows is that the context for the selection of *l*-forms of 3rd accusative clitic must include consonant-final clitics and verbs. Allomorph selection must in addition be associated with the deletion of the preceding consonant. As we shall see in Chapter 7, a morphophonological rule will be provided to capture the data both inside clitic clusters and at the verb-enclitic boundary.

4.4.3.2 Clitic fusion

Further evidence supporting the morphological status of the clitic cluster is provided by the 'fused' clusters given in (81). They comprise dative clitics (i.e., 1st/2nd/3rd Sg and 3rd pl) followed by 3rd accusative clitics and trigger vowel deletion on the preceding dative clitic.

(81) Portmanteau clusters

	3sg.masc.acc	3sg.fem.acc	3pl.masc.acc	3pl.fem.acc
1sg.dat	mo (= me+o)	ma (= me+a)	mos (= me+os)	mas (= me+as)
2sg.dat	to (= te+o)	ta (= te+a)	tos (= te+os)	tas (= te+as)
3sg.dat	lho (= lhe+o)	lha (= lhe+a)	lhos (= lhe+os)	lhas (= lhe+as)
3pl.dat	lho (= lhes+o)	lha (= lhes+a)	lhos (= lhes+os)	lhas (= lhes+as)

Although it might be argued that vowel deletion is quite common in the phrasal phonology of EP, it is worth mentioning that the forms in (81) do not result from this productive deletion rule.

Post-syntactic vowel deletion occurs in (82) where the masculine article *o* 'the.masc' is preceded by a 1st person singular dative clitic. Through vowel deletion, the determiner is incorporated into the enclitic syllable, and we obtain the 'fused' form [mu] which is homophonous with the portmanteau clitic *mo* 'dative.1sg-accusative.3.sg.masc' (compare 81 with 82):

(82) O João deu [-me o] livro. (= [mu])
the J. gave -3rd.sg.dat the.masc book
'J. gave the book to me.'

Even though the forms are homophonous, there are crucial differences between the phrasal sequence [mu] in (82) and the cluster [mu] in (81): the hiatus in (82) can be broken up through glide epenthesis, which shows that vowel deletion in (82) is optional. On the contrary, there is no alternative way of producing the clitic clusters in (81). Then there is the also the important fact that the pronoun *me* can be separated from the determiner by a pause or by word units (83), whereas nothing can break up a clitic cluster. This evidence then indicates that portmanteau clusters do not constitute phrasal combinations of two word-level units.

(83) O João deu -me ontem o livro.
the J. gave 1st.dat yesterday the.masc book
'J. gave me the book yesterday.'

In this thesis, the forms in (81) will be derived through a process of vowel deletion which affects the dative clitic. However, an alternative account might be adopted which treats the cluster as a portmanteau form. In Chapter 7, the latter view is adopted, but either account would capture the clitic-specific nature of the phenomenon.

4.4.3.4 Syncretism

One further affixal property of clitic clusters is syncretism. This morphological phenomenon take place when one affix realises two or more distinct sets of morphosyntactic features. What this means is that certain clitic combinations do not distinguish certain featural distinctions.

In the EP, syncretism is found on the following fused clusters:

(84) portmanteau clusters

	3Acc.Masc.Sg	3Acc.Fem.Sg	3Acc.Masc.Pl	3Acc.Fem.Pl
3Dat.Sg 3Dat.Pl	lho	lha	lhos	lhas

Each pair of clusters in (84) neutralises the number features of the dative clitic, reasling either a 3rd person *singular* dative clitic or a 3rd person *plural* dative clitic. So, the verb form in (85) could either mean that books were given to one person (85i) or that they were given to more than one person (85ii).

(85) Os professores deram -lhos.
the teachers gave -dat.3rd/acc.3rd.sg
i) 'the teachers gave them to him/her'
ii) 'the teachers gave them to them'

Different types of syncretism exist in inflectional systems and it is beyond the scope of this thesis to address this phenomenon in any detail (Stump 2001, Chapter 7). However, to capture the fact that these clusters are ambiguous with respect to their morphosyntactic features, one could formulate an inflectional rule that leaves the dative agreement features unspecified for number. Other derivations of syncretism which make use of rules of referral could be considered (Stump 2001, Chapter 7).

4.5 Summary

The evidence presented in this chapter lends support to the claim that clitics in EP behave morphologically like affixes. Section 4.2 shows that a wide range of data supports the claim that enclitics form a cohering unit with the verb: like verbal suffixes they a) select a verbal host, b) cannot be separated from the verb, c) cannot be coordinated, d) exhibit non-productive allomorphic variation, and e) trigger idiosyncratic stem allomorphy on the verb. Section 4.3 also shows that that clitic suffixes may intervene between the verb stem and the future/conditional marker, and section 4.4 provides evidence suggesting that clitic clusters behave like sequences of affixes.

Based on the data just presented, this thesis argues that the EP clitic system is inherently inflectional. In chapters 6 and 7 I investigate an inflectional account: chapter 6 addresses verb-final and verb-internal clitic suffixes, and chapter 7 examine the morphophonological properties of clitic-induced allomorphy. Before offering an analysis, however, an excursion into the phonology of European Portuguese is in order. Since previous phonological studies have argued against the affixal status of clitics, it seems crucial to examine the data supporting those claims to determine whether the evidence stands up to closer scrutiny.

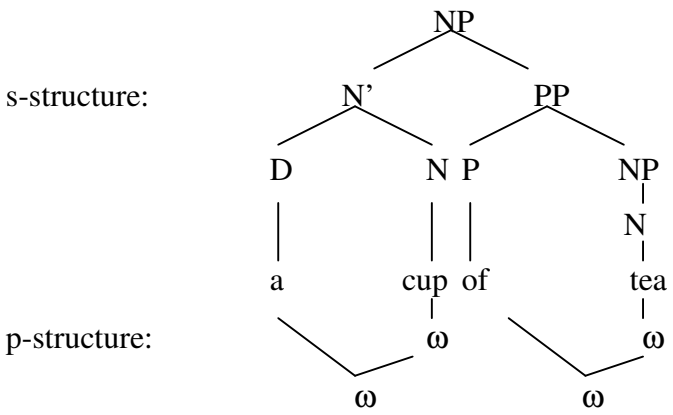
Chapter 5 Phonological evidence

This chapter argues that clitic pronouns in EP do not constitute stressless function words, contrary to recent claims by Vigário (1999a,b) within Prosodic Phonology (Nespor&Vogel 1986). To better understand the relevance of my discussion, I will start to offer a brief overview of the basic assumptions underlying prosodic approaches to cliticisation (5.1). Section 5.2 then addresses the arguments against the affixal status of clitics individually: section 5.2.1. and 5.2.2 examine the phonological behaviour of clitics drawing on a wide range of morphophonological rules and word-level rules (Luís 2001b, 2003b). Section 5.2.3 investigate the status of EP clitics as stress-neutral affixes, a property which is also attested in ordinary affixation. It is argued that the phonological properties adduced by Vigário (1999b) do not invalidate the affixal status of clitics.

5.1 Background

Within prosodic phonology (Nespor&Vogel 1986), clitic pronouns in Romance have been standardly regarded as word-level units with the prosodic properties of function words. One of the reasons for this association is the fact that both clitics and function words are typically unstressed (Zwicky 1977). This means that they, even though they constitute autonomous syntactic words in the syntax, they are word-level units without word stress. Therefore, the phonological structure of function words is not isomorphic with its syntactic structure, as in (1). Whereas content words such as *cup* and *tea* correspond both to a syntactic node and a phonological word, function words such as the determiner *a* and the preposition *of* must instead adjoin to an adjacent stressed host, i.e., the content word in (1), with which they form an 'extended' phonological word (Booij 1996, for Dutch; by Selkirk 1995, for English; by Wiese 1996, for German; by Vigário 1999a, for EP). (The symbol ω refers to the phonological word, also pwd).

(1)

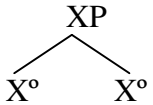


(2) [a [cup]$_{Pwd}$] $_{Pwd}$ [of [tea]$_{Pwd}$] $_{Pwd}$

Like function words, clitic pronouns cannot stand on their own in the string and require something to 'lean on'. Two basic claims are generally made about pronominal clitics: they have word-level status and attach to the host through purely phonological mechanisms (Peperkamp 1997, for Spanish and Italian clitic pronouns; Nespor&Vogel 1986). For EP, this view has been adopted by Mateus (1983), Mateus&Andrade (2000), and, specially, Vigário (1999a,b), who take the position that the verb and the clitic in EP correspond to two independent syntactic elements. The clitic-verb combination *o sentiu* in (3a) would be given the syntactic structure in (3b):

(3) a. (não) o sentiu
not 3.sg.dat felt
'didn't feel it/him'

b.



c. [o [sentiu]$_{Pwd}$] $_{Pwd}$

The representation of clitics as Xº elements, reflects the classification found in traditional grammars, such as Cunha&Cintra (1986), or Bechara (2000), in which clitics are classified as 'Pronouns' and assumed to have the syntactic status of words.

5.2 The evidence

To determine whether there are indeed arguments in favour of a phrasal analysis, as argued by Vigário (1999a,b), I will start by re-examining various phonological rules which arguably show that EP clitics behave like function words. The use of phonological rules to determine the categorial distinction between affixes and non-affixes is quite standard in prosodic phonology. It is based on the assumption that there are phonological rules in each language which apply to function words and clitics, but which fail to be triggered word-internally (Hayes 1985, Nespor&Vogel 1986). Section 5.2.1 addresses the morphological differences that are assumed to exist between affixes and clitics by Vigário (1999b). Section 5.2.2 addresses the putative phonological similarities between clitics and function words also proposed by Vigário (1999b).

5.2.1 Morphophonological rules

Among the phenomena which according to Vigário (1999a) indicate that enclitics are outside the morphological domain of the verb are a) theme vowel deletion, b) non-back glide insertion and c) theme vowel centralisation.

5.2.1.1 Theme vowel deletion

In most Romance languages, some verbal tenses do not exhibit a theme vowel (Roca 1999). In the case of EP verbal inflection, Mateus (1975) argues that the absence of the theme vowel is the result of the application of the phonological rule of theme vowel deletion. This rule, which applies

whenever a theme vowel appears before a vowel initial suffix, takes places word-internally before a person-number marker in (4) and before a tense morpheme in (5).

(4) a. *com -**e** -o → b. como
eat -Class2 -1sg.Pres.Ind
'I eat'

(5) a. *com -**e** -a -s →b. comas
eat -Class2 -Subj/Imp. -2sg
'(that you) eat'/'Eat!'

Based on this rule, it has been argued by Vigário (1999b) that if clitics are indeed affixes, vowel-initial clitics should also trigger deletion of the theme vowel. However, as the structures in (6) illustrate, when vowel-initial enclitics are preceded by a theme vowel, deletion does not take place. The fact that enclitics do not induce theme vowel deletion is taken as indication that clitics must be word-external elements (Vigário 1999b).

(6) a. com -**e** -o → b. *com-o
eat -Class2.3sg.PresInd -3sg.masc.acc
'(s/he) eats it.'

b. lava -**a** -a → b. *lav-a
eat -Class.3sg.PresInd -3sg.fem.acc
'(s/he) washes her/it.'

The first observation about Vigário's argumentation is that it overlooks two crucial aspects of the data. To begin with, it is vital to point out there are other vowel-initial suffixes, such as Preterite agreement endings, which do not trigger deletion of the theme vowel:

(7) lav -a -i [lɑvaj]
wash -TV -3sg.Pret
' (I) washed'

(8) part -i -u [pɑrtiw]
wash -TV -3sg.Pret
'(I) broke'

Instead, the suffixes *-i* and *-u* semivocalise and form a diphthong with the preceding vowel (cf. Mateus 1975, Mateus&Andrade 2000). The data in (7-8) clearly suggests that Mateus' rule of theme vowel deletion can only be triggered by specific, not just any, vowel-initial suffix. What seems to be the case, is that the rule is morphologically conditioned, as would be expected of a morphophonological process.

Let us then look at the exact environment within which the rule takes place. Mateus&Andrade (2000:75) explicitly emphasise that the theme vowel is deleted when the adjacent vowel is the tense/agreement suffix -o in the Present Indicative, and the tense-mood-aspect endings [e] and [a] (from the second and third conjugation) in the Present Subjunctive. The contexts within which the rule applies are illustrated below.

(9) Present Indicative
a. *fal(a)-o→ falo b. *bat(e)-o → bato c. *part(i)-o → parto

(10) Singular Present Subjunctive forms of *falar* 'speak', *beber* 'drink', *partir* 'break'

	Conjugation -ar	Conjugation -er	Conjugation -ir
1sg	fal(a)-e → fale	beb(e)-a → beba	part(i)-a → parta
2sg	fal(a)-e-s → fales	beb(e)-a-s → bebas	part(i)-a-s → partas
3sg	fal(a)-e → fale	beb(e)-a → beba	part(i)-a → parta

What (9-10) clearly show is that theme vowel deletion is dependent on specific tense and agreement combinations. It takes place before the 1st sg Present Indicative suffix and before the subjunctive suffixes *-e* and *-a* (cf. also Mateus et al 1989). From this we may conclude that there is no empirical evidence suggesting that vowel-initial clitics should be among the group of suffixes

triggering the phenomenon. Our first counter-argument then has shown that the affixal status of clitics cannot be determined on the basis of their ability to induce theme vowel deletion.

This brings us to our second point, namely the morphological differences between the theme-less verb forms in (4-5) and the cliticised forms in (6). What I would like to show is that both sets of verb forms have, morphologically speaking, very little in common. The crucial observation is that the vowel-final verbs in (6) - which serve as the basis for the enclitic - do not contain a theme vowel *per se*. What we need to observe is that the so-called theme vowel is in fact realising tense and agreement features, unlike the 'underlying' theme vowel in (4-5), as the glosses provided for each verb form show. The ability for certain affixes to be associated with two or more sets of featural information is a well-known property of inflectional systems (Matthews 1974, 1991). In agglutinative verb forms, one suffix realises one set of features. For example, in (11), there is an individual exponent for each one of the verbal features: *-e* for conjugation class, *-ra* for tense-mood-aspect and *-mos* for person/number.

(11) com -e -ra -mos
drink -Class2 -PluPerf -1pl
'we ate'

In (6a), however, there is only one exponent realising the values for the same set of features. In other words, the final *-e* in *come* is associated to the values a) 3rd person, b) Singular and c) Present Indicative. If we now look at (4), it is clear that the tense and agreement features are conveyed by the portmanteau suffix *-o*, not by the theme vowel (Roca 1999). The morphological structure of the data in (4-5), where deletion applies, is significantly different from the data where it is blocked. While one case contains an 'exclusive' class marker, the other doesn't.

To sum up, then, empirical evidence has been provided to argue that the rule of theme vowel deletion cannot be used to determine the word-level status of clitic pronouns in EP. Two arguments support this conclusion: a) the rule, as formulated by Mateus (1975), is only triggered by a handful

of tense and agreement suffixes in the Present and Subjunctive tenses; b) the verb forms used by Vigário to illustrate the non-application of the rule do not satisfy the requirements for the rule to apply; in particular, vowel-final verb forms which serve as the basis for enclitics, as in (6), do not contain a theme vowel *per se* and therefore fail to provide the necessary context for the rule to apply. Vigário's arguments therefore do not constitute evidence against the affix-status of clitics[28].

5.2.1.2 Non-back glide insertion

Vigário also argues against the affix status of EP clitics based on the phenomenon of non-back glide insertion. This rule is formulated by Mateus (1975) as a process which inserts a non-back glide between two adjacent vowels (provided the first vowel is a stressed [e]). The effect of the rule is illustrated with two Present Indicative forms of the verb *recear* 'fear' which show insertion of a glide between two adjacent vowels:

---

[28] The purely morphophonological character of the rule suggests that theme vowel variation should be captured as a purely allomorphic phenomenon, rather than through an analysis which assumes an underling theme vowel. The obvious advantage of an allomorphic approach is that we can do without a deletion rule which, as just shown, has a very limited context of application. For example, Roca (1999) provides empirical and theoretical arguments against postulating underlying theme vowels in Romance. To capture the idiosyncrasy of the phenomenon, it would be more insightful to assume that the morphology provides theme-less stems (Aronoff 1995) and that these stems are associated with a specific combination of tense&agreement features. One would assume, for example, that 1st singular forms of the Present Indicative select a theme-less stem which combines with the 1st singular Present Indicative marker *–o*, as in I, while other verb forms would select stems with theme vowel.

(i) 1st Singular Present Indicative of PARTIR:

Root	→ part-
Stem formation for 1sg.PresInd	→ part-
P/N agreement	→ part+o

Under such an analysis, the fact that cliticised verb forms do not trigger deletion would fall out naturally.

(12) rece **-** [j] **-** o → *receo
fear - [j] - 1sg.PresInd
'(I) fear'

(13) rece **-** [j] **-** as → *receas
fear - [j] - 2sg.PresInd
'(you) fear'

Cliticised verb forms also seem to provide the necessary phonological context for the rule to apply, as (13-15) appear to show. In these verb forms, a stressed [e] is followed by an immediately adjacent vowel, i.e. a vowel-initial enclitic. According to Vigário, if enclitics were suffixes one would expect to find an epenthetic glide breaking up the hiatus. Yet, the rule fails to apply and the hiatus between both vowels remains.

(13) dê -o → b. *de-[j]-o
give.3.sg.Imp - 3sg.masc.acc
'give it!'

(14) vê -o → b. *ve-[j]-o
see.2sg.Imp. - 3sg.masc.acc
'see it!'

(15) lê -o → b. *le-[j]-o
read.2sg.Imp - 3sg.masc.acc
'read it!'

In this respect, Vigário argues, clitics behave more like function words. For example, in (16), we have a vowel-final verb *dê* 'give.Imp' followed by a definite article *o,* but glide insertion fails to take place (cf. the similarity between 13-15 and 16).

(16) a. Dê o livro ao médico
give the book to-the doctor
'Give the book to-the doctor'

b. *Dê-[j]-o livro ao médico

There is a fundamental problem with this argument, because it presupposes that the insertion of the non-back glide takes place after *any* stressed [e] in prevocalic position. Upon closer inspection, however, the application of the rule is far from being as general as assumed by Vigário. If it did apply to any stressed *e,* one would expect to find glide insertion in (17-18), given that here the stressed [e] is also followed by a vowel-initial suffix.

Clarifying the data in (17-18), what we have are two Preterite form of *bater* 'hit' and of *comer* 'eat'. These forms contain a vowel-final stem which is produced with a stressed [e] and an agreement suffix *-u*. However no insertion of an epenthetic vowel is allowed (17-18b). Instead the *-u* suffix is semivocalised (17-18c):

(17)	a.	bate-u hit-Class2-1sg.Pret '(s/he) hit'	b. *bate + [j]+ u	c. bate-[w]
(18)	a.	come-u eat-Class2-1sg.Pret '(s/he) ate'	b. *come +[j]+ u	c. come-[w]

Crucial for an analysis of this rule seems to be the morphological status of the stressed /e/: it appears inside the root in (14), but on the theme vowel in (17-18). In Mateus (1975:147, fn.27), who formulated the rule within SPE phonology, the context of insertion of the glide is in fact described as being restricted to stressed *root* vowels [29], belonging to a small set of irregular verbs ending in –*ear*, like *recear*, *passear*. Thus, glide insertion can only apply if the vowel is part of the root.

[29] Further supporting the idea that the stressed [e] may be lexically marked is the discussion by Mateus (1975) about the arguably complex phonological status of this vowel.

Based on the observation that the rule is sensitive to the morphological status of the stressed [e], we can easily explain why glide insertion cannot apply to (15). The stressed /e/ before the enclitic constitutes a stem vowel (realising the values for conjugation class, tense, person and number), not a root vowel. We therefore conclude that cliticised verbs do not provide the necessary morphological context for the rule to apply. This rule, then, cannot be used as evidence against the affix-status of enclitics.

5.2.1.3 Theme vowel centralisation

The last rule in this group is known as theme vowel centralisation. Proposed originally by Mateus (1975), it accounts for the fact that third conjugation verbs change the theme vowel [i] into a schwa. This is illustrated in (19) with 3sg forms of the Present Indicative of *partir* 'break' and *mentir* 'lie':

(19) a. part **-e** [ə], *[i]
break -Class3.3sg.PresInd
'(s/he) leaves'

b. ment **-e** [ə], *[i]
lie -Class3.3sg.PresInd
'(s/he) lies'

Vigário (1999b) further observes that the rule fails to apply when the theme vowel is followed by other inflectional suffixes, thus suggesting that theme vowel centralisation cannot apply word-internally.

(20) a. part **-i** **-**r -e -mos
break -Class3 -Inf -Fut -1pl
'We will leave'

b. ment **-i** **-**r -e -mos
lie -Class3 -Inf -Fut -1pl
'We will lie'

Based on the data in (20), it would appear that enclitics should block the rule if they are morphologically part of the verb. This however is not borne out, as (21) illustrates.

(21) a. part **-e** -o
break -Class3.3sg.PresInd -3sg.masc.acc
'(s/he) breaks it'

b. ment **-e** -me
lie -Class3.3sg.PresInd -1sg.dat
'(s/he) lies to me'

Underlying Vigário's argument is the claim that centralisation only applies to word-final vowels. As I will show next, this claim is not accurate, for in the Present Indicative forms of *mentir* 'lie' and *partir* 'break' centralizations takes place even though the vowel is followed by an agreement marker, as shown in (22). The first problem with Vigário's claim then is that centralisation does not apply in word-final position.

(22) a. ment **-e** **-**s
lie -Class3. 2sg.PresInd -2sg.PresInd
'(you) lie'

b. part **-e** **-**s
break -Class3.2sg.PresInd -2sg.PresInd
'(you) break'

Perhaps one reason why vowel change takes place in (19) and (21-22) is the fact that the theme vowel is preceded by a stressed syllable (cf. Mateus 1975). Compare, for example, (23) with (24), where stressed vowels are given in boldfaced capitals.

(23) a. pArt **-e**
break -Class3. 2sg.PresInd
'(s/he) breaks'

b. pArt -**e** -o
break -Class3. 2sg.PresInd -3sg.masc.acc
'(s/he) breaks it'

(24) a. part -**I**a -s
break -Class3.Imperf -2sg.PresInd
'You would break'

b. part -**i** -r -Á -s
break -Class3 -Inf -Fut -2sg
'(you) will break'

This data then seems to show that the rule is determined by stress properties of the verb form. This insight is also formalised in Mateus (1975) within SPE-Phonology. This means that the factors derterming the vowel change are not related to word-final boundaries, thus predicting that the presence of an enclitic does not have any effect on how the theme vowel of third conjugation stems surfaces.

5.2.1.4 Summary

This section re-examined morphophonological rules previously used as evidence against the affixal status of EP clitic (Vigário 1999b). Upon closer inspection of the data, the following conclusion were drawn:

- Theme vowel deletion (5.1.1.1) and non-back glide insertion (5.1.1.2) select lexically and morphologically conditioned contexts that are not available at the verb-enclitic boundary. Failure of the rule to apply to enclitics is simply the result of the highly specific context of application of these rules.
- The rule of theme vowel centralisation (5.1.2.3) derives verbal stems for third conjugation lexemes; the rule is determined by the verb's stress pattern and is therefore not affected by the presence or absence of enclitics.

Based on these conclusions, I conclude that none of the above morphophonological rules provides evidence against the morphological status of clitics.

5.2.2 Word-level rules

In this section, I address another set of rules that have been used to support the claim that clitics behave phonologically like function words: a) back vowel deletion (cf. 5.1.2.1), b) nasal glide insertion (cf. 5.1.2.2), c) high vowel semivocalisation (cf. 5.1.2.3) and d) non-back vowel deletion (cf. 5.1.2.4). As in the previous section, it will be argued that the arguments do not stand up to closer scrutiny.

5.2.2.1 Back vowel deletion

Let us start with the rule of back vowel deletion, proposed by Frota (1996), which may optionally delete a back vowel in word final position when followed by another vowel. It applies between two prosodic words (i.e. stressed content words), as in (26), but fails to be triggered if function words are involved (27).

(26) músic**o** africano → músic∅]$_{Pwd}$ africano]$_{Pwd}$
'african musician'

(27) d**o** architecto → d*∅]$_{Fnc}$ architecto]$_{Pwd}$
of-the architect

The fact that the rule is also blocked by clitics has been interpreted as an indication that enclitics and function words are phonologically similar (Vigário 1999a).

(28) eu não t**o** aceito → eu não t*∅ aceito
I not 2sg.dat/3sg.masc-acc accept
'I don't accept it from you'

Yet, the fact that the rule of back-vowel deletion also fails to apply across morpheme boundaries clearly shows that this claim cannot be upheld. For instance in the verbs forms *voar* 'fly' and *doar* 'donate' of the first conjugation, (29), the root vowel –o is followed by the theme vowel –a, but deletion of the first vowel cannot take place:

(29) a. d**o**ar → d*∅ar
'donate'
b. v**o**ar → v*∅ar
'fly'

Instead, the data suggests that back-vowel deletion is not only blocked in the syntax, as in (27), but also word-internally, as in (29). If this observation is correct, then the rule cannot be used to examine whether clitics behave like affixes or like function words because it seems to identify only adjacency between prosodic words, as in (26) (Frota 1996). Failure of the rule to apply between the verb and the enclitic does not constitute evidence against the affix-status of clitics.

5.2.2.2 Nasal glide insertion

The rule of nasal glide insertion, as proposed by Mateus (1977), changes the nasal vowel [ẽ] into the nasal diphthong [ɐ̃j̃] [30]. It reflects a standard assumption in Portuguese phonology that the diphthong [ɐ̃j̃] can only take place word-finally (Bechara 2000, among other), as the monomorphemic words in (30a) and the morphologically complex words in (30b) seem to show and as the verb forms in (31) illustrate.

(30) a. hom**em**]$_{Pwd}$ *[ɔmẽ], [ɔmɐ̃j̃]

[30] The insertion of a nasal glide after the nasal vowel [ẽ] triggers centralisation of the vowel, in the following context: [ẽ] →[ɐ̃j̃] __ [j̃]

‘man’

b. virag**em**]$_{Pwd}$ *[viraʒẽ] , [viraʒɑ̃j]

‘turn (n.)’

(31) a. faz**em**]$_{Pwd}$ *****[fazẽ], [fazɑ̃j]

‘(they) do/make’

b. diz**em**]$_{Pwd}$ *[dizẽ],[dizɑ̃j]

(they) say

On the contrary, the underlying nasal vowel [ẽ], which is in complementary distribution with the diphthong, occurs word-internally, across morpheme boundaries (32) or inside underived lexemes (33):

(32) a. **en**tendemos [ẽtẽdemuš] , *[ɑ̃jtẽdemuš]

‘(we) understand

b. **en**fiamos [ẽfjɑmuš] , *[ɑ̃jfjɑmuš]

‘(we) insert’

(33) a. m**en**tol [mẽtɔl] **,** *****[mɑ̃jtɔl]

‘mint’

b. qu**en**te [kẽt] , *****[kɑ̃jt]

‘hot’

The assumption that the nasal diphthong [ɑ̃j] cannot be followed by a morpheme boundary is used to argue that clitics cannot be morphologically part of the verb. Note that enclitic pronouns can follow nasal glides (cf. 34). The fact that clitics fail to block diphthongisation therefore indicates that their attachment to the verb must take place in phrasal phonology, as would be expected of function words.

(34) a. diz**em** -lhe *[dizẽ-ʎə], [dizɑ̃j-ʎə]

say.3sg.PresInd -3sg.dat

‘they say to them’

b. faz**em** -te *[fazẽ-tə], [fazɑ̃j-tə]
do.3pl.PresInd -2sg.acc
'they do to you'

There is however an important piece of evidence which clearly contradicts the view that the nasal vowels only appear in word-final position. As discussed in Chapter 4, there is the crucial fact that when the nasal ending is part of 3pl verb forms, it induces phonological shape variation on 3rd accusative pronouns, as shown in (35). As argued in Chapter 4 (cf. section 4.3), morphophonological data of this nature can only be derived as a morphophonological phenomenon because it does not follow from general rules of productive phonology.

(35) a. *dizem -**o** → dizem-**no**
say.3pl.PresInd -3sg.mas.acc
'they say it'
b. *fazem -**as** → fazem-**nas**
do.3pl.PresInd -3pl.fem.acc
'they do'

In particular, *n*-initial clitic allomorphs (i.e., *-na, -no, -nas, -nos*) are not simply triggered by the adjacency of any nasal diphthong. If they were, one would expect the 3sg verb form *mantém* to induce allomorphy, contrary to the evidence.

(36) *mantém-**na** → mantém-**a**
keep.3sg.Pres.Ind -3sg.fem.acc
's/he keeps it/her'

Pronominal allomorphy is instead grammatically conditioned, that is, it is determined by specific person and number features of the verb form they combine with. This evidence constitutes one of the strongest arguments in favour of a morphological account of cliticisation which assumes that enclitics combines with the verb in the morphology (cf. Spencer 1992, Crysmann 1997, Luís 2001b,

2003a). I therefore conclude that nasal glide insertion is not restricted to word-final position, for in verb forms they trigger idiosyncratic phonological change on enclitics (cf. section 5.3.2.2 for informal analysis). This means that there is no evidence weakening the affixal status of enclitics.

In addition, it is important to observe that Vigário adopts a purely phonological approach to the derivation of the diphthong *-em*, arguing that it is underlyingly a nasal vowel. From a morphological point of view, however, the verb forms in (34) contain the diphthong simply because they realise the same agreement ending which in EP is -em, for the Present Indicative of the second and third conjugations. Under this alternative account, diphthongisation fails to provide any support in favour of Vigário's claims.

5.2.2.3 High vowel semivocalisation

The rule of high vowel semivocalisation is found in the Lisbon dialect of EP and applies to vowels in prevocalic position (Mateus 1975). It takes place before a masculine gender suffix in (37), and before a verbal suffix in (38).

(37) a. ri**o** *[ʀi**u**]/[ʀi**w**]
'river'

b. ti**o** *[ti**u**]/[ti**w**]
'uncle'

(38) a. sorr-i-**o** *[suʀi**u**]/[suʀi**w**]
smile-Class3-1sg.Pres.Ind
'I smile'

b. ri**o** ***[ʀi**u**]/[ʀi**w**]
laugh-Class3-1sg.Pres.Ind
'I laugh'

In (39), where the high vowel appears word-finally, semivocalisation is blocked, indicating that the rule cannot apply across word boundaries, and that its domain of application is word-internal.

(39) a. Não v**i**]$_{Pwd}$ **u**tilidAde]$_{Pwd}$ nisso *[**iw**] , [**iu**]
not saw usefulness in-that
'I didn't see any usefulness in that'

b. Eu vi **o**]$_{Pwd}$ carro]$_{Pwd}$ *[**iw**] , [**iu**]
I saw the car
'I saw the car'

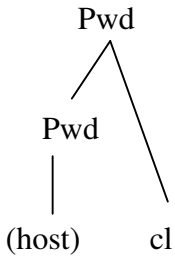
For the present discussion, what is interesting about this rule is that it also takes place at the boundary between the verb and the enclitic, as in (40), showing that the rule treats enclitics and suffixes as identical elements.

(40) a. Ela v**i** **-o** [v**iw**]/ *[v**iu**]
she saw.1sg -3sg.masc.acc
'She saw-him/it'

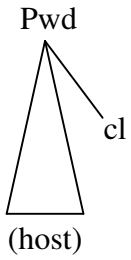
b. Eu fing**i** **-o** [fĩʒ**iw**]/ *[fĩʒ**iu**]
I pretended.1sg -3sg.masc.acc
'I pretended it'

For Vigário (1999a), this similarity is rather unwelcome and an account is provided which attempts to explain why enclitics and suffixes behave alike. Vigário tries to argue that this similarity takes place because there is no prosodic boundary between the verb and enclitics. To remove the boundary, Vigário makes use of the distinction between 'prosodic adjunction' and 'prosodic incorporation', common in Prosodic Phonology (Nespor&Vogel 1986). Both adjunction and incorporation take place in phrasal phonology, combining phonological words with unstressed function words:

(41) a. adjunction



b. incorporation



While the phonological attachment in (41a) creates a prosodic boundary between the phonological word and the function word, the one in (41b) unites both the function word and the prosodic word under one phonological word domain without internal boundaries (Peperkamp 1997). Each one of these patterns of prosodic attachment are based on phonological evidence which arguably shows that adjunction and incorporation trigger different phonological phenomena. Underlying these different types of prosodic attachment is the claim that certain rules of phonology are sensitive to prosodic boundaries, being either blocked or triggered by them.

Based on this theoretical distinction, Vigário (199a) argues that enclitics are rather special function words in that they attach to their host through incorporation, unlike standard function words (e.g., determiners or prepositions) which attach through adjunction. For EP, then, the fact that high vowel semivocalisation is blocked by prosodic word boundaries, as in (37-38), motivates an incorporation process which combines the enclitics with the verb without introducing any phrase-phonological boundary (Vigário 1999a). In this respect, enclitics behave phonologically like suffixes, not like function words, because it is also assumed that no prosodic boundary exists between verbal stems and verbal suffixes.

Let us now address the next rule, before commenting on Vigário's proposal.

5.2.2.4 Non-back vowel deletion

Phrasal incorporation for enclitics is also proposed on the basis of the rule of non-back vowel deletion (Vigário 1998, 1999b). This rule deletes vowels in prevocalic position before a phonological word boundary (Frota 1996).

(42) ped**e** azeitonas → ped∅]$_{Pwd}$ azeitonas]$_{Pwd}$
ask.3sg.PresInd olives
'(s/he) asks for olives'

The rule is blocked *between* the verb and an enclitic (43), but not *after* the verb-enclitic unit (44). This suggests that cliticised verbs do not contain any internal phonological boundary (45c).

(43) a. ped**e** -a → b. *ped∅ - a
3sg.ask.Pres.Ind -3sg.fem.acc
'(s/he) asks for it/her'

(44) a. dou-t**e** amêndoins → b. dou-t∅ amêndoins
give-2sg.dat peanuts
'(I) give you peanuts'

(45) a. ped**e**]$_{Pwd}$ azeitonas]$_{Pwd}$ (cf. 42)
b. ped**e**-a]$_{Pwd}$ (cf. 43)
c. dou-t**e**]$_{Pwd}$ animais]$_{Pwd}$ (cf. 44)

Under the assumption that the domain of application of the rule is the right edge of a phonological word (Frota 1996), the position is taken that enclitics must incorporate into the phonological domain of the verb in phrasal phonology (Vigário 1999a), as suggested for the rule of high vowel semivocalisation addressed in 5.2.2.3.

There are, then, two pieces of evidence which seem to motivate the process of (phrasal) prosodic incorporation for enclitics, assigning enclitics a marked status as function words. Recall that in Vigário's work, function words are by default adjoined to a prosodic word.

In what follows, however, I will try to show that neither of the rules does support Vigário's claim about prosodic incorporation. Instead, these rules severely undermine the view that enclitics are phonologically different from suffixes, showing that enclitics behave suffixes because they *are* suffixes.

Even though Vigário uses the evidence in (42-45) to motivate the status of enclitics as incorporated function words, the same data could well be used to argue in favour of affixal status. If the rule is blocked by word-internal boundaries, then treating enclitics as verbal suffixes equally prevents the rule from applying. So, the claim that there is no phrasal boundary between the verb and the enclitic is clearly compatible with the affixal view of enclitics. Not only is non-back vowel deletion blocked before an enclitic, but it is also blocked word-internally, between other non-back vowels, in this case a root vowel and the following theme vowel (46).

(46) rec**e** **-**a -va → *rece∅ - va
fear -Class1 -Impf.1/3sg
(I, s/he) feared

In addition, the similarity between affixes and enclitics is also supported by the fact that non-back vowels can also be replaced by glides (Vigário 1998, 1999a). This replacement however is also sensitive to the affixal and non-affixal status of the word: while glide insertion is ungrammatical (or more marked) across prosodic words (47a), it is mandatory before enclitics and word-internally (47b-c). This fact seems to suggests that the rule is treating enclitics and suffixes identically.

(47) a. ped**e** azeitonas → */?ped[j] azeitonas (cf.42)
b. ped**e**-a → ped[j]-a (cf. 43)
c. rec**e**ava → rec[j]ava (cf. 46)

To sum up, then, accounting for the data in 5.2.2.4 through an unmotivated process of phonological incorporation completely fails to do justice to the fact that enclitics and verbal suffixes

exhibit similar behaviour. If this observation is correct, then this rule would in effect substantially weaken the claim that enclitics are function words. The rule therefore then cannot be used to support the non-affixal status of clitics, and therefore the claim that enclitics are special function words is not tenable.

In addition, the rule of semivocalisation (5.2.2.3) has also traditionally been regarded as a word-internal rule (Mateus 1975). The fact that it applies between the verb and the enclitic strongly suggests that postverbal clitics and verbal suffixes share indeed identical phonological properties, severely undermineing the view that enclitics are phonologically different from suffixes[31].

5.2.2.5 Summary

This section has shown that there is insufficient evidence supporting the claim that EP clitics behave like function words. In particular, the rule of back-vowel deletion (5.1.2.1) proved inconclusive because it applies both to function words *and* affixes. The remaining rules, namely nasal glide insertion (cf. 5.1.2.2), high vowel semivocalisation (cf. 5.1.2.3) and non-back vowel deletion (cf. 5.1.2.4) fail to support Vigário's claim and support instead affixal status of enclitics.

An independent discussion of Vigário (1999a,b) is provided in Crysmann (2002). Unlike Crysmann (2002), however, I have provided empirical arguments against the so-called word-level rules in 5.2.2 which either apply to ordinary affixes (in which case, the phenomena are morphologically triggered and do not constitute evidence in favour of a postlexical derivation) or they simply do not apply under the conditions defined by Vigário (1999a,b).

[31] In section 5.3, it is argued that the phrasal incorporation of enclitics suggested by Vigário (1999a) poses serious problems to the behaviour of enclitics with respect to stress assignment.

5.3 Stress

In this section, I shall consider one further aspect of the phonological behaviour of EP clitics, namely their inability to affect stress placement. Within phonology, this property has been used to weaken an affixal analysis because of the tendency for verbal suffixes in Romance to trigger migration of stress (Harris1994, Gerlach 2001a, Vigário 1999a,b). Section 5.3.1 briefly surveys previous phonological accounts of stress and examines how they have been applied to cliticised verbs. Section 5.3.2 suggests an alternative account based on the simple assumption that clitic pronouns in EP, and also in Spanish and Italian, constitute *stress-neutral* affixes (Monachesi 1999). Based on evidence from languages with stress-neutral affixes, it will be argued that stress-neutrality should not constitute an argument against the inflectional approach to the data. I shall then attempt to accommodate the stress-related behaviour of enclitics within an inflectional analysis of cliticisation.

.

5.3.1 Overview

One well-known phonological property of EP and Romance verbal suffixes is the fact that they induce rightward migration of word stress (Roca 1999, Mateus&Andrade 2000). In (48), stressed vowels are given in bold face and syllable boundaries are signalled with "-".

(48) a. l**a-**va
'(s/he) washes'

b. la-v**a-**va
'(I) washed'

c. la-va-r**i-**as
'(you.sg) would wash'

On the contrary, postverbal clitics behave quite differently with respect to stress placement. As (49) illustrates, stress falls on the theme vowel regardless of whether the verb is followed by a postverbal clitic or not.

(49) a. can-t**a-**mos
'we sing'
b. can-t**a-**mos-lhe
'we sing for him/her'

Stress migration, as illustrated in (48), ensures that stress falls in one of the three last last syllables of the verb. Because enclitic suffixes are stress-neutral, stress on cliticised verbs may fall outside the 3-syllable window. In (50), stress falls on the fourth to last syllable. The inability for clitics to affect the placement of verbal stress is also found in Italian and Spanish. In (51b), stress appears on the fifth to last syllable.

(50) a. can-t**á**-va-mos
'we sang'
b. can-t**á**-va-mos-lhe
'we sang for him/her'

(51) a. m**é-**sco-la
'stir'
b. m**é-**sco-la-me-la (Italian, Monachesi 1999)
'stir it for me'

Based on such evidence, it has been claimed that stress-neutral clitics cannot be regarded as part of verbal morphology because they do not satisfy the necessary stress-related requirements for affixhood (Vigário 1999a,b, for EP; Harris 1994, for Spanish, Peperkamp 1997 for Italian).

In the following sections, I will provide an overview of current approaches to stress placement in clitic structures which will serve to argue that stress neutrality does not constitute a strong argument against an inflectional approach.

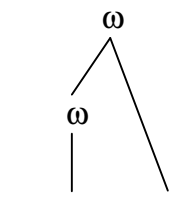
5.3.1.1 Previous accounts

Under the standard assumption that the locus of word stress is the phonological word (represented as ω), a syntactic terminal element is typically regarded as the domain within which stress rules apply cyclically[32].

(52)

As to clitics, it is argued that because they have no effect on the locus of stress they are not part of the ω domain of stress assignment. This is represented as in (53) where the lower ω constitutes the stress domain, also known as the innermost phonological word. Under this view, stress neutrality is captured by placing clitics outside the phonological domain of the verbal host (Booij 1995, for Dutch, Peperkamp 1997, for Italian and Spanish).

(53) (Neapolitan, Peperkamp 1997)



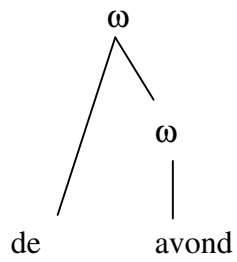
'tell.Imp it'

In (53), the clitic adjoins in phrasal phonology to a lower ω forming with it a higher ω (cf. 5.1). In this example, the upper ω corresponds to a phrasal construction, while the lower ω corresponds to an inflected verb form. Under this view, clitics are assumed to behave phonologically like function

[32] It is standardly assumed that the phonological word can be smaller or larger than the terminal syntactic node or equal in size (Booij 1983). The combination of a function word with its stressed host is assumed to constitute a phonological word corresponding to two syntactic nodes.

words. As (54) shows, the Dutch function word *de* 'the' is also adjoined to a lower ω which constitutes its prosodic host.

(54) (Dutch, Booij 1996)



'the evening'

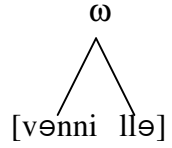
It is worth mentioning that there are also stress-shifting clitic pronouns in Romance. Evidence has been provided by Peperkamp (1997) and Monachesi (1999) for Italian dialects such as Lucanian and Napoletano. In Lucanian, for example, clitic pronouns force main stress to shift onto the penultimate:

(55) /vín:e le/ → [vən:íl:ə] (Lucanian, Peperkamp 1997)

'sell.Imp it'

Surprisingly, when clitics do affect word stress prosodic studies do not adopt an inflectional view of cliticisation but instead argue that these clitics behave prosodically in an expectional way. In (56), it is assumed that the phonological word corresponding to the verb *vənni* can freely restructure in phrasal phonology and incorporate enclitics. Through phrasal incorporation (not adjunction) it is assumed that a phonological word such as 'vənni'is extended (recall that incorporation has also been adduced by Vigário (1999a) for EP enclitics to account for two phonological rules; cf. discussion in 5.2.2.3-4 and 5.3.1.2)

(56)



'sell.Imp it'

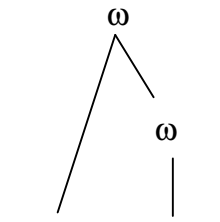
Given this representation, clitics are placed inside the inner phonological word, as part of the domain where stress rules apply (Peperkamp 1997, Booij 1996). Rather than adjoining to ω as in (53), they incorporate into it. Note, however, that the cliticised verb in (56) corresponds syntactically to two syntactic nodes: one for the verb and another for the so-called function word. Incorporation is a postlexical (or phrasal) process.

5.3.1.2 The case of EP

We have seen just seen that stress-shifting clitics are assumed to be located inside the (innermost) phonological word, while stress-neutral clitics are assumed to be outside this domain. Crucially, both adjunction and incorporation are assumed to take place in phrasal phonology. As we shall see now, the proposals made for EP clitics diverge somewhat from the 'orthodox' view within Prosodic Phonology.

The idea that clitics must be outside the domain of stress assignment is also adopted by Vigário (1999a) within Prosodic Phonology. The representation of proclitics is given below:

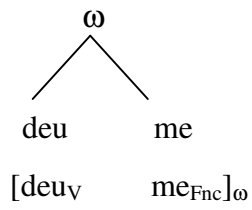
(57) (Vigário 1999a)



'(that you) speak to me'

For enclitics, however, the phonological structure is different.

(58)



So, as alluded to earlier (cf. 5.2.2.3-5.2.2.4), while proclitics attach through adjunction, enclitics attach through incorporation. The distinct prosodisation is motivated by the rules of high vowel semivocalisation and non-back vowel deletion.

As far as stress placement, the problem with the representation in (58) is that it predicts that EP enclitics, like Lucanian enclitics, affect stress placement[33]. One might argue that stress placement is switched off before cliticisation takes place or alternatively it might be argued that stress placement rules are restricted to the domain of the 'inner prosodic word'. Both these proposals however would be purely stipulative and would clearly contradict the assumption shared among phonologists that stress rules apply cyclically within the phonological word[34]. Incorporation is therefore in significant departure with the assumption that stress-neutral clitics are placed outside the domain of stress assignment through the process of adjunction (Peperkamp 1997, Booij 1996).

Unlike Vigário (1999a), Mateus (1983) correctly positions enclitics outside the domain os stress assignment. However, unlike Vigário, she does not deal with the behaviour of enclitics with respect to the rules of high vowel semivocalisation and non-back vowel deletion addressed in 5.2.2.3-5.2.2.4.

---

[33] This suggestion is also made in Vigário (1999a) who agrees that this may be a potential problem for her analysis.

[34] Mateus' (1983) phonological analysis of stress does not encounter this problem because enclitics are simply adjoined postlexically, as the word-boundary in (67), indicates.

In short, then, previous analyses of the stress behaviour of EP enclitics suffer from either empirical or theoretical shortcomings. What we need therefore is an account which accommodates the stress neutral status of EP clitics within the affix-like behaviour (evidenced in section 5.2).

5.3.2 Stress-preserving clitics

While stress-neutral affixes are quite abundant cross-linguistically, function words triggering stress migration are clearly more marked. In addition, enclitics, as shown before, trigger other morphophonological effects (cf. Chapter 7) which clearly indicate that they should be regarded as affixes.

It is well known in morphology that both inflectional and derivational affixes can be either stress-preserving or stress-shifting. For example, the derivational suffixes *–ness*, *-y* and *-ous* in English leave the word stress of their base unaffected, as illustrated in (59), where stress falls on the same syllable both on the underived and the derived forms. As a result, stress may fall outside the three-syllable constraint, as in (60), where stress occurs on the fourth to last syllable (Aronoff&Shridhar 1983, Aronoff 1976, Hammond 1999).

(59) a. h**A**ppy
b. h**A**ppi-ness
(60) a. **a**ccura(c)-y
b. sp**I**ritu-al

Similar data can be found in Dutch, where derivational affixes are also classified as either stress-neutral or stress-preserving (Booij 1995).

Likewise, inflectional affixes do not necessarily shift word stress. Regular affixes in English, such as *–ing*, and the Italian 3rd pl suffix *-no* of the Present Indicative and Conjunctive paradigm can

appear on verbs with stress on the fourth to last syllable (Nespor&Vogel 1986, Monachesi 1999, Peperkamp 1997):

(61) part**i**cipating

(62) a. tel**é**fonano
'they phone' (Pres.Indic)

a'. tel**é**fonino
'(that) they phone' (Conjunctive)

b. c**ó**mplicano
'they complicate' (Pres.Indic)

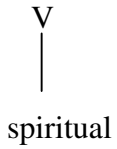
b'. c**ó**mplichino
'(that) they complicate' (Conjunctive)

Pre-antepenultimate stress in Italian is regarded as a violation of the 3-syllable window referred to above (i.e, that stress must occur on one of the three last syllables of the word). However, even though it is stress-neutral, there appears to be little evidence suggesting that the Italian suffix *–no* is not morphologically part of the verb. In fact, a syntactic analysis of this agreement marker would fail to show that it behaves in all other respects like a verbal suffix (e.g., it cannot be separated form the verb, only appears postverbally and stands in paradigmatic opposition with other agreement endings). With such cross-linguistic evidence, it is clear that the ability to induce stress-shift cannot be regarded as a necessary condition for affix-hood.

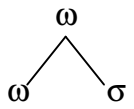
It is not only empirical evidence that supports the idea that affixal clitics can be stress-neutral. Conceptually, the need to assign theoretical status to stress-neutral affixes is quite consensual within phonology. For example, Level Ordered Morphology, which proposes a distinction between root/stem-affixes and word-affixes, assumes that word-affixes do not affect lexical stress of the base because they are added to the base after lexical stress has been assigned (Wiese 1996). The same intuition is captured within Optimality Theory by Hammond (1999) who proposes that stress-neutral affixes must be outside the 'inner' morphological domain within which stress assignment is operative. Under this approach, stress-neutral affixes are simply regarded as *marked* morphological elements which incur a violation of the constraint which determines that all affixes must be inside the domain of stress-assignment.

Common to these approaches is the assumption that the domain within which stress assignment is operative can be smaller than a terminal syntactic node, or to put it differently, it can constitute the morphological basis to which stress neutral affixes attach. This may be schematically represented as follows:

(63) a. syntax



b. (word-level) phonology



[spiritu al]$_A$

The adjective *spiritual*, which constitutes a terminal syntactic node in (63a), is mapped onto the derived base *spiritu-,* namely the innermost phonological word (ω) to which stress applies. The stress-neutral suffix *-al* is then adjoined to the lower ω, so as to capture the fact that it is stress-neutral. Crucially, the upper ω which dominates *spiritual* is derived exclusively in the morphology[35].

Summing up, various arguments have been provided to support the view that the property of stress-neutrality should not constitute an argument against the affixal-status of clitics. Instead, if one takes the wide range of morphological evidence as primary, then the inability to shift word stress is simply accounted for as the result of an incomplete grammaticalisation process, without putting into question the morphological nature of the data. This is precisely what we will assume for EP clitics.

---

[35] Here both the lower and upper ω are contained within the morphological word (cf. 5.3.1 for examples where the lower ω is derived in the morphology and the upper ω in phrasal phonology.

5.3.2.1 Clitic suffixes

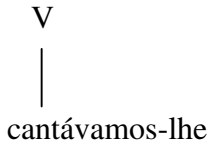
Following Monachesi (1999) on Italian, I will now show that EP clitic pronouns can be analysed as stress-neutral affixes (Luís 2001b, 2003b).

a) enclitics

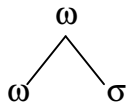
Let us start with enclisis and examine the cliticised verb fom *cantávamos-lhe* which is stressed on the pre-antepenultimate syllable. I will assume that the verb form is mapped onto a phonological structure with an embedded phonological word. This means that the verbal base *cantávamos* is mapped onto the lower ω and constitutes the domain within which stress applies; the stress-neutral clitic, on the contrary, is placed outside the inner prosodic word in (65c).

(65) a. cantávamos-lhe
'(we) sang for him/her'.

b. Syntax



c. Word-internal phonology



[cantávamos lhe]$_V$

Recall that the phonological structure in (65c) is being assigned to a genuine morphological word. The structure is therefore not phrasal.

The proposal in (65) nicely predicts stress-neutrality, because the fact that enclitics are outside the innermost ω predicts that enclitics are not available for stress assignment (cf. Monachesi 1999, for Italian). This proposal also enables us to make correct predictions about double stress in 'mesoclitic' verb forms.

b) Mesoclisis

Recapitulating somewhat, in Chapter 4 it was mentioned that enclitics in EP appear not only verb finally but also verb-internally between the stem and the tense/agreement endings of the future and conditional forms. One of the problems posed by these verb forms is that they bear two stresses: one before the clitic (i.e., in the theme vowel) and another one after the clitic: in the latter case, stress falls on the vowel or diphthong immediately preceded by the clitic. This stressed vowel or diphthong can appear in various morphological contexts: a) followed by person/number suffixes (66a), b) followed by another tense vowel (66b), or c) merged with person/number features (68).

(66) a. lav**a**r-te-**e**mos
wash-2sg.acc-fut.2pl
'we will wash you'

b. lav**a**r-te-**í**amos
wash-2sg.acc-cond.2pl
'we would wash you'

Under the assumption that words have only one main stress, mesoclitic verb forms have been analysed by some as periphrastic/analytic constructions made of a three word-level elements (cf. Chapter 4): the verb, the enclitic and the future/conditional agreement marker (## and # signal word boundaries; + introduces morpheme boundaries):

(67) ## bat**e** + r +# te # **á** + s ## (adapted from Mateus 1983)

Under this view, double stress is derived by treating the tense ending *-ás* and the verbal base *bater* in (67) as word-level units that can be assigned lexical stress independently. Ordinary verb forms, as given in (68), are classified as synthetic verbs with one lexical stress.

(68) *cantará* ('you will sing') (Pereira 1999)

x

x x x

cantar**á**

Even though the analysis in (67) is faithful to the diachronic origins of the future and conditional endings, there are strong arguments showing that the 'periphrastic' view of mesoclitic verbs is not tenable and that these verb forms are best derived as synthetic forms [36]. As alluded to earlier in Chapter 4, the tense endings in mesoclitic verbs are phonologically exactly identical to those in non-cliticised verb forms, as the underlining in (69) illustrates, suggesting that they constitute the same morphological elements.

(69)	a. conditional	b. (mesoclitic) conditional
	eu lavar**ia**	lav**a**r-lhe-**i**a
	tu lavar**ia**s	lav**a**r-lhe-**i**as
	ele lavar**ia**	lav**a**r-lhe-**i**a
	nós lavar**í**amos	lav**a**r-lhe-**í**amos
	vós lavar**í**eis	lav**a**r-lhe-**í**eis
	eles lavar**ia**m	lav**a**r-lhe-**i**am

Furthermore, in both (69a) and (69b) tense and agreement markers appear on the right periphery of the verb, they cannot coordinate, and, crucially, they carry word stress. Given these similarities, it seems to be more intuitive to assign the same grammatical status to both putatively distinct sets of tense suffixes. In addition, the affix-like status of clitics further weakens the periphrastic analysis, for, from a morphological viewpoint, enclitics (in final or medial position) are part of verbal

[36] On the diachronic evolution of mesoclisis see Vázquez Cuesta & Mendes da Luz (1971), Roca (1999), Pereira (1999), Leeuw (1997).

morphology[37]. From this viewpoint, the verb forms in (69b) should be derived as inflected verb forms.

Based on these arguments, the question now is how to derive the co-existence of two stresses: one on the stem, another one on the tense marker. In this study, the answer to the problem will be based on two straightforward assumptions. First, we adopt the claim presented in the previous section that the domain of lexical stress assignment does not include the enclitic. The stem of *lavar-te-ei*, namely *lavar-*, forms the domain of stress assignment, while enclitics as well as future/conditional endings are outside this domain. In addition, I also assume that Future and Conditional tense endings are inherently stressed suffixes. Since stress on the tense suffix is not derived through phonological rule, we predict that stress bearing suffixes, such as the bold faced endings in (69b), can co-occur with the stressed stem. It is also important to observe that the intervening unstressed clitic prevents both stresses (i.e. the one on the stem and the one on the tense suffix) from being adjacent to each other, thus further enabling their co-occurrence.

To summarise, then, the stress properties of mesoclitic verb forms can be captured if we assume that:

(70) Assumption 1:
The morphological basis of enclitics (i.e., the clitic host) coinstitutes the domain of stress assignment.

(71) Assumption 2:
Future and Conditional markers are inherently stressed suffixes (in both the cliticised and the non-cliticised paradigm).

In analogy with our analysis of enclitics, the domain within which lexical stress is assigned is represented as the inner ω. This constituent, which is at the same time a stress domain *and* a

[37] A more detailed discussion of the morphological status of mesoclitic verb forms is given in Chapter 6. See also Chapter 4 for further details on the problems with the 'analytic' view of mesoclitic verb forms.

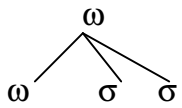
morphological base, combines with the enclitic and the inherently stressed tense/agreement marker to form one and only one syntactic terminal node:

(72) a. lavar-lhes-ia
'(I) would wash for them'

b. Syntax

lav**a**r-lhes-**i**a

c. Word-internal phonology



[lav**a**r]$_{\omega}$ -lhes -**i**a]$_{\omega}$

Viewing future/conditional suffixes as inherently stressed suffixes is not just mere stipulation. It is crucial to observe that ordinary future and conditional verb forms, as in (69), also display an idiosyncratic stress pattern (Roca 1999). Unlike other tenses, in which stress falls generally on either the theme vowel or the root (cf. Mateus 1983, Mateus et al 1989, Pereira 1999), in future and conditional forms it falls on the tense marker. This is briefly illustrated in (73) with Present Indicative forms and in (74) with future forms.

(73) a. l**a**v -a -s
wash -Conj.Cl1.Pres.Ind -2sg
'(you) wash'

b. lav -**a** **-**mos
wash -Conj.Cl1.Pres.Ind -2sg
'we wash'

(74)

eu lavar**ei**	'I will wash'
tu lavar**á**s	'you.sg will wash'
ele lavar**á**	'(s)he will wash'
nós lavar**e**mos	'we will wash'
vós lavar**ei**s	'you.pl will wash'
eles lavar**ão**	'they will wash'

Indeed, stress never falls on the root or on the theme vowel in Future and Conditional verbs but rather on the vowel or diphthong placed immediately after the -r suffix of the 'infinitival' stem[38]. Future and Conditional tenses exhibit the morphological structure given in (75), where TV stands for theme vowel; 'r' stands for the so-called infinitival suffix which combines with vowel-suffixes to convey future and conditional tense; and P/N corresponds to person and number agreement.

(75) morphological structure of future and conditional verb forms in EP:

Root -TV - r - Fut/Cond- P/N

As is typical in Romance, the theme vowel can be either -a, - e, or –i. Future and Conditional tense can be realised as a separate exponent, as in (76a,b), or in combination with P/N, as in (76 c,d,e). In either case, we find the following two types of future and conditional markers: a) 'r'+ vowel or 'r'+diphthong (in the future, cf. 76c-d) ; b) 'r'+ two vowel suffixes (in the conditional cf. 76e).

[38] The so-called infinitival stems is a genuine example of a parasitic stem (Mattews 1972, Aronoff 1994). Although it is phonologically exactly identical to a true infinitival verb forms, it serves as the stem for the derivation of a finite tense (cf., ch. 6 for further discussion).

(76) a. lav -a -r -ía -mos
root -TV -r -Cond -P/N
'(we) would wash'

b. beb -e -r -á -s
root -TV -r -Fut -P/N
'(you.sg) will drink'

c. part -i -r -á
root -TV -r -Fut.P/N

d. com -e -r -ei
root -TV -r -Fut.P/N

e. lav -a -r -ia
root -TV -r -Cond.P/N
'(I/(s)he) would wash'

The reason why future and conditional suffixes attract lexical stress has been attributed to the historical development of these verb forms. The Romance future developed from the periphrastic construction *infinitive+habere*.Pres.Ind while the Romance conditional derives from *infinitive+habere*.Imperf (cf. Harris&Vincent (1988), Leeuw (1997), Roca (1999) for historical development). Assuming that both the infinitive verb and the auxiliary were stressed, it has been argued that when the construction evolved into a synthetic form only one of them could be preserved, in this case it was the rightmost stress.

As Roca (1999) and Leeuw (1997) observe, the fact that Future and Conditional violate the general rule should indicate that the idiosyncratic stress pattern is determined by a lexical property of the Future/Conditional vowel-suffixes. Under this view, stress on the Future/Conditional suffixes is derived from their lexical representation rather than from a specially conditioned phonological rule. This view also extends to the derivation of the 'synthetic' paradigm in (69a) and (74), where it is assumed that the general phonological rule which assigns stress to the theme vowel is overruled by the morphological stress rule (Roca 1999).

5.3.2.2 Phonological rules revisited

The assumption that clitics should be regarded as stress-neutral affixes also allows us to account for the phonological rule of nasal glide insertion addressed in section 5.2.2(Luís 2003b).

As alluded to earlier, nasal glide insertion applies either word-finally or word-internally (between the verb stem and an enclitic suffix). The nasal diphthong appears at the end of the verb in (77a) and word-internally in (77b) (allomorphic alternation of the enclitic has taken place). Glide insertion however fails to apply in any other word-internal contexts (78).

(77) a. diz**em** [dizɐ̃j̃]

b. diz**em**-no [dizɐ̃j̃-nu]

'(they) say it'

(78) a. **en**fiar *[ẽfjaʀ]

'to insert'

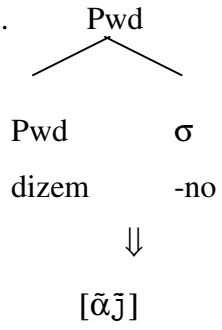
b. qu**en**te [kẽt]

'hot'

To account for the glide insertion in (77b), it may be argued that the application of the rule is triggered by the prosodic boundary which the enclitic introduces (cf. 79), while the word form in (80) lacks such a boundary.

(79) a. diz**em** no '(they) say it'

b.



(80) a. enfiar 'insert'

b. Pwd

enfiar

⇓

[ẽ]

If we accept with Mateus (1977) that nasal glide insertion should be analysed as a phonological rule[39], then the rule might look as follows:

(81) [ẽ] → [ɐ̃j̃] / __]ω

(82) diz**em** [ẽ] → [ɐ̃j̃]

The above rule simply says that a nasal vowel is diphthongised before a phonological word boundary, which can be either word-internal or word-final. This rule then accounts for the occurrence of the nasal glide word-internally, as in (79), but also for its word-final placement in (82).

5.4 Summary

Underlying most prosodic approaches to cliticisation is the assumption that clitic pronouns have the same prosodic status as function words: both are argued to be prosodically deficient and attached (either through adjunction or incorporation) in phrasal phonology. This claim has been made for EP by Vigário (1999a,b), within Prosodic Phonology, following the work of Mateus 1983 (cf. also Mateus&Andrade 2000). In her work on EP, Vigário (1999b) tries to support the similarity between

[39] Under a morphological point of view, it would be simply assumed that the diphthong constitutes a 3rd plural agreement marker which is not sensitive to the presence of enclitics (cf. 5.2.2.2).

object pronouns and function words by using rules which set affixes and clitics far apart. This chapter however argues that there is insufficient empirical support to uphold her claim. In particular, in section 5.2, numerous phenomena were examined which showed that enclitics and function words do not share all the phonological similarities assumed by Vigário. The affixal status of EP clitics has so far not been weakened.

One phonological property of EP clitics which appears to weaken their morphological status is the fact that they do not trigger stress shift, unlike 'genuine' verbal sufixes in Romance. This aspect was addressed in section 5.3. It was suggested that clitics can be represented as stress-neutral affixes, outside the domain of stress assignment. Further arguments against the view that the verb-enclitic unit should be derived in phrasal phonology are discussed in Chapter 7 on the basis of non-productive allomorphy found at the boundary between verbs and enclitics.

Summing up the thesis so far: Chapters 4 and 5 have examined the morphological behaviour of EP clitics by providing extensive affixal data and phonological evidence. The next chapter proceeds with a preliminary inflectional analysis of the clitic suffixes. Other aspects of clitic suffixation, in particular the morphophonology, will be postponed until Chapter 7.

Chapter 6 Clitic suffixes

This chapter explores an inflectional account of clitic suffixes in EP within Paradigm-Function Morphology (Stump 1993, 2001), adopting insights from Spencer (2000) and Spencer&Luís (in press). Special attention is given to both verb-final clitics, as found in *perguntavas-me* 'you asked me', and verb-internal clitics, as in *perguntar-me-ás* 'you will ask me'.

Section 6.1 investigates how cliticised verb forms can be derived as purely inflected words, examining several hypotheses. The fact that clitic suffixes in EP attach either after all ordinary inflections, like in other Romance languages, or before the future and conditional tense marker raises interesting questions about the type of morphological host they can attach to.

The discussion about the inflectional analysis of clitic suffixes then leads us to consider the verbal paradigm of Portuguese. A sketch of how Portuguese ordinary suffixation can be derived within PFM is provided in section 6.2. Section 6.3 finally extends the inflectional analysis to cliticisation. Since clitic suffixes can appear inside the verb form, it will be suggested that inflectional rules realising clitics must be subject to a reordering algorithm which enables clitic suffixes to attach to an infinitival stem. By default, however, clitic suffixes attach to a fully derived verb form.

6.1 General issues

The data provided in Chapters 4 and 5 strongly suggest that postverbal clitics in EP should be analysed as part of verbal morphology:

(1) Affix-like properties of suffixing clitics in EP
   a) combine with verbs in strict adjacency;
   b) induce allomorphy on the verb and undergo allomorphic variation;
   c) appear word internally before Future and Conditional tense markers;

d) display distributional, semantic and morphophonological idiosyncrasies inside the clitic cluster;
e) behave phonologically like word-internal elements.

(For similar data on Romance clitics see Anderson 1996**,** Auger 1994, Brines 2001, Legendre 2000a, Miller 1992, Miller&Sag 1997, Monachesi 1999, Crysmann 2002).

As this thesis will show, even though clitics behave *very much* like affixes (as opposed to word-level units), there is a sense in which all morphologists agaree that clitics are not *exactly* like ordinary affixes. Any inflectional analysis must therefore be able to capture the fundamentally inflectional behaviour of clitics but at the same time show that clitics constitute an extended type of inflection. Affixal clitics, in this respect, constitute an interesting challenege to linguistic theory. This point will be made clearer as we go along in the discussion, specially when we address the format of paradigm-functions, i.e., the inflectional rules determining the inflected forms for a given lexeme (cf. Chapter 2). These rules, which in Paradigm-Function Morphology define ordinary verbal paradigms Stump (2001), will be extended in order to capture the idea that cliticised verbs are part of the extended verbal paradigm of EP.

6.1.1 Identity of the host

This section starts by examining the morphological base of clitic affixes. In the literature it is widely assumed that pronominal clitics in Romance languages such as Italian and French attach to the edge of fully formed verbs (Anderson 2000, Miller&Sag 1997, Monachesi 1999, amongst other). In this section, however, I shall argue that even though this assumption appears to be correct for the majority of the Romance languages, it fails to account for the verb-internal placement of clitics in EP.

6.1.1.1 The standard view

Perhaps one of the most distinctive properties of ordinary affixes is their need for a morphological base. For ordinary affixation, the host generally constitutes an underived base, either a root or stem (Matthews 1991). This base serves as input to the inflectional rule realising a given affix. In simplified terms, we can think of the relation between an affix and its base as in (2), where the 'input' specifies the morphological host of a given suffix, while the 'output' constitutes the product of combining the base with the suffix:

(2)

	input	Output
Rule 'suf$_1$'	*stem*	*stem*+ *suf*$_1$

When words comprise two or more affixes, it is further assumed that each inflectional rule applies to the output of the previous rule. This cyclic derivation is diagrammatically illustrated in (3) and shall be illustrated in more detail throughout this chapter on the basis of EP verbal morphology.

(3)

	input	Output
a. Rule 'suf$_1$'	*stem*	*stem*+ *suf*$_1$
b. Rule 'suf$_2$'	*stem*+*suf*$_1$	*stem*+ *suf*$_1$+*suf*$_2$

In (3), we start off with a stem and combine it with suf1 (in 3a); we then take the sequence 'stem+suf1' and combine it with 'suf2' deriving the sequence 'stem+suf1+suf2'. Thus, in realisational models of morphology (Anderson 1992) inflected words are derived through the cyclic attachment of affixes. Note in addition, that the morphological base which serves as input to the second rule in (3b) is morphologically more complex than the base which serves as input to the first one. Each base however constitutes an incomplete (i.e., underived) expression.

Let us now see how the dependency between an affix and its base is captured within the theory of Paradigm-Function Morphology (Stump 2001). The intuition is captured by defining inflectional

rules over an incomplete (i.e., not yet fully formed) verbal host, represented as 'X' in (4). This variable constitutes the more basic expression which serves as input to the rule. The right hand side of the rule provides the output of applying the Realisation Rules (RR) to the base (see also Chapter 2), as both Portuguese RRs given in (4) illustrate (cf. section 6.3.2).

(4) Realisation rules for EP (fragment) (cf. Appendix A)

a. $RR_{Ia,\ \{\ TENSE:\ Imperfect\ \}Conj.1,\ V}$ (<X, σ>) $=_{def}$ <X-*va*, σ>

b. $RR_{IId,\ \{\ AGR:\ 1\ Pl\ \},\ V}$ (<X, σ>) $=_{def}$ <X-*mos*, σ>

These are rules which realise a given set of morphosyntactic features through the suffixation of specific exponents: in (4a) the exponent is the Imperfect marker *–va* of first conjugation forms and in (4b) it is the agreement marker *–mos*. Both rules participate in the derivation of the 2pl Imperfect form *levávamos* ('we took') of the lexeme LEVAR, applying sequentially to the verbal stem *leva-*: we start the derivation of *levávamos* 'we took' by applying RRI_a to the stem *leva-* of LEVAR 'carry' yielding the underived form *levava-*; we then apply the RR_{IId} with the underived verb form *levava-* to yield the complete verb form *levávamos* (see section 6.2 for details).

Thus, two important properties of ordinary affixes are: a) they are dependent on a morphological host and b) they construct inflected words from more basic (underived) expressions. Let us now see how affixal clitics are assumed to attach to their morphological host.

When comparing the canonical attachment of ordinary affixes with that of pronominal clitics, one cannot fail to observe that clitics constitute affixes of a different sort. In particular, in most Romance languages pronominal clitics appear after all verbal suffixes, as the italicised forms show, suggesting that unlike ordinary affixes these affixal clitics constitute external layers of morphology:

(5) a. Tu *chantes*. (Fr.)
'You are singing'

b. *Chantes-tu*?
'Are you singing'

(6) a. Quiero *comprar* libros. (Sp.)
want buy books
'I want to buy books'

b. Quiero *comprarlos*.
want buy-acc.3sg
'I want to buy them'

To account for the peripheral placement of clitics in Romance, both in preverbal and postverbal position, various proposals have been made in the literature. It is generally proposed that clitics combine with fully inflected verb forms. Nishida (1987), within Categorial Grammar, for example, formulates cliticisation as a rule of verb formation which maps between V_1 and V_2, where V_1 is a fully inflected verb from, e.g. *comprar* '(to) buy', and V_2 is a cliticised verb, e.g. *comprarlo* '(to) buy it'. Similarly, Monachesi (1999, 2000) within HPSG determines that the 'stem' of a clitic is, in fact, an inflected word; likewise, Miller&Sag (1997) derive cliticised verbs in French by taking inflected forms (called 'I-form') as the input to pronominal affixation. In all these studies it is assumed that a word form (which could appear as an independent word form in the syntax) can also be the morphological host of a clitic pronoun[40].

This intuition has also been formulated within the theory of Paradigm Function Morphology, by Spencer (2000), for Macedonian clitic pronouns. One of the questions addressed in that study is how to formulate inflectional rules so as to capture the fact that clitics attach to fully formed verb forms. For the derivation of cliticised verb forms in Macedonian, Spencer (2000) adopts the standard format for RRs (cf. 4). However, instead of defining the morphological base as an underived base, Spencer defines it as a fully inflected verb form, represented for convenience as 'V' in (7).

[40] This is view is different form Anderson (1992, 1996, 2000) where clitics are not part of verbal morphology. We return to Anderson's treatment of phrasal affixation in Chapter 8 (cf. also Chapter 3).

(7) RRs for Macedonian pronominal clitics (Spencer 2000)

a. RR $_{I,\ \{CASE:\ Dat,\ PERS:\ 1sg\},\ V}$ (<V, σ>) $=_{def}$ <*mi*-V>

b. RR $_{II,\ \{CASE:\ Acc,\ PERS:\ 3sg,\ GEN:fem\},\ V}$ (<V, σ>) $=_{def}$ <*ja*-V>

As given in (7), clitics constitute the phonological expression of a morphosyntactic feature set. Under this view, the RR in (7a) says that the dative 1st person singular clitic *mi-* realises the feature set {CASE: Dat, PERS: 1sg} and the rule in (7b) says that accusative 3rd person singular feminine clitic *ja-* constitutes an exponent of the feature set {CASE: Acc, PERS:3sg, GEN:Fem}. In addition, (7) further informs that clitic-RRs (i.e., RRs deriving clitics as affixes) take a fully formed verb as input and derive a cliticised verb as output.

In Spencer's analysis, it is assumed that an inflected verb form such as *donesete* 'bring.Imperative' can combine sequentially with the clitic suffixes *–mi* and*–ja* to derive the cliticised verb form *donesete mi ja* 'bring me it'. The cliticised verb woud be derived as follows: the rule in (7a) takes as input the imperative verb form *donesete* yielding the form *donesete mi*; and (7b) applies to the output of (7a) and yields *donesete mi ja*. The RRs in (7) then have the important effect of extending an inflected verb form by adding new layers of affixation, and inflectional rules are assigned the ability to apply to a morphologically complete base.

Summing up, then, it is generally assumed that rules realising ordinary affixes are defined over an underived base (i.e., stem or root), while rules realising affixal clitics apply over inflected words yielding 'extended' verb forms. So, in contrast with (2-3), cliticisation can be diagrammatically represented as in (8) where the inflectional rule realising a pronominal clitic takes as input a complete verb form.

(8)

	Input	Output
Rule 'cl_1'	*verb form*	*verb form*+cl_1

6.1.1.2 The problem

What this section will show now is that the inflectional approach suggested by Spencer (2000) for Macedonian and by other linguists for French, Spanish and Italian cannot be adopted for EP clitic suffixes. Even though enclitics in EP occur - as in all other Romance languages - at the right edge of a fully formed verb, i.e., after all ordinary verbal inflections, they interact with ordinary tense and agreement markers in ways which suggest that they must be defined over an underived base like ordinary suffixes.

In (9), the Imperfect form *levávamos* forms the base for the clitic suffix *-te* in (9b). For this type of data, the claim that inflectional rules realising clitics are defined over a V, as in Spencer (2000), is correct.

(9) a. Levávamos as crianças à escola.
took.2sg.Impf the children to school
'we took the children to school'

b. Levávamos-te à escola
took.2sg.Impf-acc.2sg to school
'we took you to school'

However, an alternative view seems to be necessary if we want to make the right predictions for EP mesoclitics. Here I take on a remark by Spencer (2000:384, fn. 12) who points out that clitics can sometimes interact with verbal affixes in such a way that 'outer' layers appear as 'inner' layers. EP provides a good example of this type of interaction with the phenomenon of 'mesoclisis'. In (10b) and (11b), clitics occur verb-internally because the verb is in the future tense (10b) and in the conditional tense (11b)[41]:

[41] Even though the forms in (10c) and (11c) are ungrammatical in standard EP, they are attested in non-standard varieties. This synchronic co-occurrence of two forms of cliticised future and

(10) a. falarás
speak.Fut.2sg
'(you) will speak'

b. falar-nos-ás
speak-acc.2pl-Fut.2sg
'(you.sg) will speak to us'

c. * falarás-nos
speak.Fut.2sg-acc.2pl
'(you.sg) will speak to us'

(11) a. falarias
speak.Cond.2sg
'(you) would speak'

b. falar-me-ias
speak-acc.2pl-Cond.2sg
'(you) would speak to me'

c. *falarias-me
speak.Cond.2sg-acc.2pl
'(you) would speak to me'

Various arguments support the view that (10b) and (11b) constitute morphologically cohering verb forms: on the one hand, the nature of tense and agreement markers, on the other, the behaviour of the 'mesoclitic' (cf. Chapters 4 and 5). As to tense suffixes, they are formally and distributionally exactly identical to the endings on non-cliticised future and conditional verbs; likewise the regular agreement endings are the same throughout the whole EP verbal conjugation. With regard to the clitic suffix, it displays all the affixal properties assigned to clitic suffixes in verb-final position, such as inducing non-productive allomorphy and combining into clitic clusters.

---

conditional verb forms seems to suggest that there are two grammars in contemporary EP: the homogenous grammar (in which enclitics are always verb-final) and the grammar in which clitic placement is sensitive to tense features.

The question then is how to capture the fact that clitic suffixes appear verb-internally. While for the example in (11b) it could be simply assumed that clitics are peripheral suffixes, the same cannot be said about the examples in (10b) and (11b). Mesoclisis therefore requires an alternative account of clitic attachment.

Hypothesis 1

One way of accounting for clitic placement in verb-medial position would be to propose the hypothesis in (12):

(12) "enclitics, but not mesoclitics, take as input an inflected verb form"

From a purely empirical point of view, this hypothesis would combine enclitics with fully formed verbs and attach mesoclitics before tense and agreement endings. Conceptually, however, the attachment of mesoclitics (but not enclitics) to the stem, would entail that there are two types of realisation rules: mesoclitics would be realised through realisation rules similar to those adopted for ordinary suffixes (cf. 4), while enclitics would be generated through typical clitic-like rules (cf. 7). Mesoclitics would thus would be more affix-like than enclitics.

This analysis however would fial to explain why enclisis is more productive than mesoclisis, why enclitics and mesoclitics are formally exactly identical, why they display the same inventory of clitics, combine into the same set of clitic-clusters, and, crucially, trigger and undergo the same range of morphonological alternations (cf. Chapter 4). There is then no evidence suggesting that enclitics and mesoclitics should be regarded as different suffixes. Any account must be able to show that verb-internal and verb-final clitic suffixes constitute one and the same clitic suffix.

Hypothesis 2

Assuming that there can be only one set of inflectional rules realising clitic suffixes, we need to decide whether they are both realised through the same set of clitic-rules (combining with V) or

through the same set of ordinary rules (combining with some underived base). I will argue below that only the latter hypothesis is tenable. First, however, let us see why the rules in (7) are not feasible.

Even though the rules in (7) would correctly derive verb-final placement, they would not be able to derive mesoclisis in an insightful way. Given that these rules would apply to fully formed verbs, we would be forced to say that the clitic suffix in *levar-lhe-emos* 'we will bring them' attaches to the fully formed infinitival verb *levar* '(to) take'. Even though this is what syntactic accounts generally propose, there are several problems with this view:

a) The first one is related to the fact that there is in effect no 'infinitival' verb form in *levar-lhe-emos*; what we have instead is an infinitival stem which is homophonous with the verb but which does not carry any non-finite features (cf. Matthews 1972, Aronoff 1994). Note that it would make little sense to say that it did, because that would create a clash between the non-finite properties of the verbal base and the finite properties of the future/conditional markers[42]. So, the infinitival stem which appears in 'mesoclitic' verb forms is therefore, using the designation of Matthews (1972), a parasitic stem. This is an issue I shall return to in section 6.2 where I address the inflectional derivation of future and conditional verb forms.

b) One further problem with the idea that mesoclitics attach to an infinitival verb is that the future and conditional tense and agreement markers could no longer be analysed as genuine verbal suffixes. There is however little empirical support for a proliferation of endings because these putatively distinct sets of tense and agreement suffixes are morphologically absolutely identical to ordinary markers (see Chapter 4, section 4.3 for problems with viewing these markers as syntactic elements). In addition, since it is not only the future and conditional marker but also the markers for

---

[42] This problem is also found in analyses which do not regard the tense and agreement markers as verbal suffixes, cf. Leeuw (1997), Duarte et al. (1995), Vigário (1999b). See section 6.3.4 for overview of previous accounts.

person and number that appear after the mesoclitic (e.g. *lavar-lhe-e-mos* 'we will wash'), this option would entail that there are also two sets of agreement markers, further triggering the proliferation of homophonous verbal suffixes.

Given these arguments, it will be assumed that clitic suffixes (both verb-final and verb-internal) shall be realised through ordinary realisation rules, of the type presented in (4), which take as input an underived base.

6.1.1.3 Summary

This introductory section examined the nature of the morphological host which serves as the base for affixal clitics. Evidence was provided which showed that the ability for EP clitic suffixes to appear as inner layers of the verb form constitutes a challenge to the idea that clitic suffixes attach to fully formed verb. The standard distinction between ordinary affixes and clitics makes wrong predictions for EP, because in this language suffix clitics interact with tense and agreement affixes in a way that suggests that they interleave with internal layers of suffixation. Thus, any inflectional analysis of clitic suffixes in EP must be able to capture the two following points:

(13) a) enclitics and mesoclitics constitute positional variants of the same suffixal clitic;
b) mesoclitics attach to an infinitival *stem* and are followed by genuine tense and agreement endings.

In the remainder of this chapter, I shall attempt to develop an inflectional analysis to capture the insights in (13). It shall be assumed that clitic suffixes in EP take as input an underived base (i.e., a verbal stem) and that the order in which clitic suffixes and ordinary affixes appear is subject to reordering, adopting previous work on reversed affixes by (Stump 1993). Anticipating the analysis, I shall clitic suffixes undergo a reversal process which assign clitics the ability to appear either verb-finally or verb-internally (6.4). This proposal is sketched informally in (14a) for verb-final clitic suffixes and in (14b) for verb-internal clitic suffixes:

(14) a. Imperfect = stem + tense + agr + *clitic* = leva+va+mos+*lhe*

b. Future Indicative = stem + *clitic* + tense + agr = levar+*lhe*+e+mos

6.1.2 Affix reordering

Affix reordering takes place when a given sequence of affixes is reversed. To illustrate the phenomenon, let us consider the case of Fula subject/object agreement markers (a branch of Niger-Congo) and survey the analysis provided by Stump (1993, 2001) within Paradigm-Function Morphology. The phenomenon is also found in Quechua and Swazi, and is therefore not unknown to inflectional morphology (cf. also Noyer 1994).

As (15) shows, object agreement follows subject agreement in Fula verb forms. Therefore the default order of affixal positions of a Fula verb forms is as given in (16).

(15) mball -u -(no) -daa -be (Stump 1993:165)
help -Rel:Past:Act -(Pret) -you.Sg -them:CL1
'you (sg.) (had) helped them'

(16) Fula relative past tense verb form *mball-u-no-daa-be* 'you had helped them'

i. Rule block:	X	I	II	III	IV
ii. Affix:	root	u	no	daa	be
iii. Features:	help	TENSE.VOICE	PRETERITE	SUBJECT	OBJECT

In certain person and number combinations, namely when a verb form is both inflected for first person singular subject agreement and 2nd or 3rd person singular object agreement, subject and object agreement suffixes appear in the opposite order:

(17) mball -u -(no) -maa -mi (Stump 1993:165)
help -Rel:Past:Act -(Pret) -you.Sg -1
'I (had) helped you (sg)

The reversed order is illustrated diagrammatically in (18).

(18) Fula relative past tense verb form *mball-u-no-maa-mi* 'I had helped you'

i.	Rule block:	RRI	RRII	RRIV	RRIII
ii.	Affix:	u	no	maa	mi
iii.	Features:	TENSE.VOICE	PRETERITE	OBJECT	SUBJECT

Even though affix-metathesis is not an unusual phenomenon, previous theories of inflectional morphology are not very good at deriving these data. An insightful discussion is found in Stump (1993) who illustrates the difficulties of accounting for Fula morpheme reordering within a model of morphology which is not paradigm-based, such as an A-Morphous Morphology (Anderson 1992). In Stump (1993), this phenomenon is defined as a case of position class reversibility : "members of position class A precede those of position class B in one set of forms, while members of class B precede those of class A in a complementary set of forms" (Stump 1993:165).

Though the definition may look trivial, the phenomenon poses problems to some realizational models of inflection such as Andersons's A-Morphous Morphology (1992), based on linearly ordered rule blocks. As pointed out by Stump, there is no linear ordering of blocks which could generate the data in (18). The only possible way of deriving the data, without violating the spirit of linear ordering, would involve assigning the 'reversed' subject and object makers to two different rule blocks, thus increasing by two the number of affixal positions in Fula verb forms (cf. Stump 1993, for discussion). One of the problems with this analysis however is that it suggests that there are two different sets of subject and object markers, failing to capture that it is instead the same set of affixes which can appear in reversed position.

In PFM, however, affixes which undergo metathesis do not have to be assigned to different rule blocks. Instead, Stump (1993) assumes that there is an unmarked affix order in Fula which can be reversed under specific morphosyntactic conditions. To account for the Fula data, then, Stump assumes that there is a default order for the realisation of rule blocks which generates the sequence of affixes TENSE »VOICE »PRETERITE »SUBJECT » OBJECT, shown in (15-16). It is further assumed

that this sequence is overridden whenever the two following conditions are met: the value for tense is 'relative past' and the values for object agreement is 2sg or 3sg/class 1. With this combination of tense and agreement features, RRIV applies before RRIII, that is, the object marker appears before the subject marker.

The exact morphosyntactic conditions determining the reversed application of RRs are defined in Stump's (1993) Paradigm-Function analysis. Leaving aside the technical details, one of the crucial claims in Stumps' analysis is that PFs have the ability to reverse the order in which rule blocks apply. The simplicity with which PFM derives the phenomenon stems from the fact that an affix's membership to a given rule block is independent of the order in which that rule block applies and is therefore independent of the position that affix occupies in the string of affixes. In other words, although the numeral subscript assigned to a rule block reflects the order in which these blocks apply, rule block membership is defined by RRs whereas rule block application is defined by paradigm-functions (cf. Chapter 2). This shows that the interplay between RRs and PFs can elegantly capture the phenomenon of affix metathesis (cf. Stump 1993, 2001 for analysis). This intuition shall be explored in section 6.3 with respect to mesoclisis.

6.2 Portuguese verbal morphology

In preparation for the PFM-analysis of enclisis and mesoclisis to be developed in this chapter, a simple survey of Portuguese verbal conjugation will now be provided. Section 6.2.1 addresses the inflectional structure of some Portuguese verb forms; section 6.2.2 illustrates how verbal morphology is derived within the theory of Paradigm-Function Morphology (Stump 2001).

6.2.1 Description

This section describes 1st conjugation forms of the Imperfect, Future and Conditional paradigms. The fragment of Portuguese verb forms is given in Appendix A. A more complete description of Portuguese verb forms will be left for future research.

6.2.2.1 Affixal positions

To start with, it will be assumed that verb forms in European Portuguese have two affix positions: one position for tense suffixes and another position for agreement suffixes. Based on the Imperfect, Future and Conditional paradigms, the template for Portuguese verbs is as follows[43]:

(19) template for EP verb forms:
stem - tense - agreement

A concrete illustration of these position is given in (20) with 1st plural forms of *levar* 'take' from the Imperfect, Future and Conditional paradigms (cf. Appendix A for other forms). It is important to observe that (19) does not entail that all verb forms exhibit two distinct tense and agreement affixes. In fact, in 1st and 3rd Sg Imperfect and Conditional forms, and also in 3rd SG Future forms, there is no agreement marker (see Chapter 2 on the Identity Function Default for the treatment of zero exponents in realisational morphology).

(20) a. leva -va -mos
take.1st -Impf -1pl
' we take'
b. levar -e -mos
take.1st -Fut -1pl

[43] This template assumes that stems are not derived morphololgically (see below for discussion about the 'morphomic' status of stems).

'we will take'

c. levar -ia -mos
take.1st -Cond -1pl
'we would take'

These affix positions would also apply to most of the other paradigms of the first conjugation. However, for simplicity, I will only address the fragment given in Appendix A.

The tense suffixes that shall be assumed are as follows: two Imperfect suffixes, *-va* (the default) and *–ve* (for 2Pl forms). Likewise, for future tense, there is the default marker *-a* and also the marker for 1st Sg/Pl and 2nd Pl forms *-e*. In addition, there is one Conditional suffix, namely *–ia*, and *–ie* for 2nd Pl forms. For each of the tense markers, it might also be argued that the more specific variants are derived as phonologically conditioned allomorphs, assuming that palatal vowels trigger centralisation of the *-a* vowel. In this case, the Imperfect marker *–ve* and the Conditional marker *–ie* would be triggered by the 2nd Pl agreement suffix *–is*, even though this would not apply to 1pl future marker *–e*. I shall leave this question open and assume, for simplicity, that these endings are proper suffixes, not allomorphs (see realisation rules in 6.2.2).

The agreement affixes are 2nd person singular -s, 1st person plural *-mos*, 2nd person plural *-is* and 3rd person plural *-N* [+nasal]. They appear in all three paradigms (and effectively in all other paradigms, except in the Preterite). More irrgular endings are the 1st person singular marker *-i* which appears in the Future paradigm.

6.2.2.2 Verbal stems

The segmentation given in (19) is based on the assumption that only suffixes with morphosyntactic value shall be derived through realisation rules. This means that theme vowels, which constitute meaningless forms ('morphomes' in Aronoff's terminology), are not exponents of realisation rules, but are instead derived though 'morphomic' rules which make no reference to morphosyntactic

properties. They serve to define a class of stems to which some set of realisation rules systematically applies (cf. Stump 2001, Luís&Spencer forthcoming).

The distinction between inflections and morphomes is based on the widely-held claim that realisation rules derive inflections as the phonological output of morphosyntactic features, whereas in the absence of such features realisation rules cannot apply (Aronoff 1994). Instead, 'meaningless' affixes help derive stems and do not convey any 'meaning'. I shall follow Aronoff (1994) in assuming that a given language can have a rich inventory of stem formation rules and subsequently of stems. For the verb forms given in (20) I am assuming that any lexeme of EP will need, at least, the two following stem types:

(21) a. $stem_1$ = root + theme vowel -a, -e, -i (e.g., *lev*+a)
b. $stem_2$ = $stem_1$ + -r (e.g., *leva*+r, except for *dir-*, *far-*, *trar-*)

While $stem_1$ combines the root with a theme vowel, $stem_2$ combines the basic stem with the *-r* suffix (cf. also Villalva 1994 on stem selection for EP). One property of stems is that they recur in different paradigms. So, $stem_1$ is the default stem of the Present, Preterite, Imperfect paradigms. $Stem_2$ in (21b), on the other hand, is selected by the Future and Conditional paradigms. As to future and conditional verb forms in EP, it is important to observe that both sub-paradigms select the same stem, namely the 'infinitival' stem', as the table in (22) shows.

(22) Infinitive, future and conditional paradigms of COMPRAR 'buy'

Infinitive	Future	Conditional
Comprar	*comprar*-á	*comprar*-ia
	comprar-ás	*comprar*-ias
	comprar-á	*comprar*-ia
	comprar-emos	*comprar*-íamos
	comprar-eis	*comprar*-íeis
	comprar-ão	*comprar*-iam

As mentioned in Chpater 4, the diachronic reasons are well known: Romance future and conditional verb forms derive the Latin periphrastic combination of an infinitive verb with a Present or Imperfect auxiliary. In present-day Portuguese, these verb forms are fully synthetic and have therefore lost their 'infinitival' meaning. They constitute instead verb forms with finite tense. The infinitive sequence then is best regarded as an 'infinitival' stem, and as already observed, it merely provides the inflectional base for the attachment of future and conditional tense suffixes. Unlike the genuine infinitive verb, it does not correspond morphosyntactically to a non-finite form. So, to account for the case of EP future and conditional verb forms, I make use of an 'infinitival' stem (i.e., $stem_2$) which is homophonous with the infinitive verb and which will serves as the input for tense and agreement. For regular future form, then, we will need the stem type given in (23) which serves as the basis to which the future tense marker applies:

(23) levar + é + mos
$stem_2$ + aff + aff

On the other hand, irregular future and conditional stems will be indexed separately and override the default stem. So for the DIZER and FAZER, for example, the future stem is overridden by 'dir-' and far-', respectively.

(24) a. FAZER 'do': fazemos '1st pl Pres', but *far*emos/*farí*amos '1st pl Fut/Cond'
b. DIZER 'say': dizemos '1st pl Pres', but *dir*emos/*dirí*amos '1st pl Fut/Cond'

The stems *far-* in (22a') and *dir-* in (22b´), which occur in the Future and Conditional, constitute the inflectional basis to which the tense and agreement endings attach. It is also this irregular future/conditional stem which serves as the basis for cliticisation, as in *dir-lhe-emos*.

6.2.2 Analysis

I shall now offer a treatment of the EP verbal paradigm within the theory of Paradigm-Function Morphology (cf. Chapter 2). Special attention is given to the interaction between realisation rules (RRs) and paradigm functions (PFs).

6.2.2.1 Rule blocks and Realisation rules

Based on the two affixal slots suggested for the verbal template in (19), I propose the two following Rule Blocks: block I for the realisation of tense markers, and Block II for the realisation of agreement suffixes.

(25) Fragment of realisation rules (RRs):

a. Block I

i. $RR_{Ia,\ \{Tense:Impf\},\ Conj:1,\ V}$ (<X, σ>) $=_{def}$ <X-*va*,σ>

ii. $RR_{Ib,\ \{Tense:Impf,\ Agr:\ 2PL\},\ Conj\ 1\ V}$ (<X, σ>) $=_{def}$ <X-*ve*,σ>

iii. $RR_{Ic,\ \{Tense:Future\},\ V}$ (<X, σ>) $=_{def}$ <X-*a*,σ>

iv. $RR_{Id,\ \{Tense:Future\ Agr:\ 1Sg/PL,\ 2PL\ \},\ V}$ (<X, σ>) $=_{def}$ <X-*e*,σ>

v. $RR_{Ie\ \{Tense:Cond\},\ V}$ (<X, σ>) $=_{def}$ <X-*ia*,σ>

vi. $RR_{If\ \{Tense:Cond,\ Agr:\ 2PL\},\ V}$ (<X, σ>) $=_{def}$ <X-*ie*,σ>

b. Block II

i. $RR_{IIa,\ \{2sg\},\ V}$ (<X, σ>) $=_{def}$ <X-*s*,σ>

ii. $RR_{IIb,\ \{1pl\},\ V}$ (<X, σ>) $=_{def}$ <X-*mos*,σ>

iii. $RR_{IIc,\ \{2pl\ \},\ V}$ (<X, σ>) $=_{def}$ <X-*is*,σ>

iv. $RR_{IId,\ \{3pl\ \},\ V}$ (<X, σ>) $=_{def}$ <X-*N*,σ>

v. $RR_{IIe,\ \{Tense:Future\ Agr:\ 1\ Sg\ \},\ V}$ (<X, σ>) $=_{def}$ <X-*i*,σ>

In determining the number of rule blocks, I follow Anderson (1992) and Stump (2001) in assuming that, by default, each affixal slot is realized by a separate rule block. I also assume that rule blocks can only apply once during the derivation of a given word form (cf. Chapter 2).

As to the format of the RRs in (25), it has been alluded to before that each realization rule in PFM contains three indices: a) Roman capital numbers identify the rule block to which a RR belongs[44]; b) the feature set realized by each RR given within curly brackets and c) the category of the lexeme whose inflected word is being derived and which is represented as V. In addition, the pairing <X, σ> stands for the input to the rule while <X-*i*,σ> represents the output of applying the rule to X. This then means that each rule realises a given inflectional affix as the phonological realisation of a specific set of morphosyntactic features. So, the suffix *–va* is realised as the default exponent of Imperfect tense (for first conjugation forms) and the suffix *–mos* is realised as the exponent of 1st plural agreement. More accurately, we can also say that RRIa realises the feature {TENSE: Impf} through the suffixation of *–va*, and that RRIIb realises the feature set {Agr: 2PL} through the suffixation of *-mos* (cf. Chapter 2 for more detailed overview of the formalism).

6.2.2.2 Paradigm Function

Whereas RRs are individual operations which express subsets of morphosyntactic features, PFs determine how the word forms of a given lexeme are inflected (cf. Chapter 2). The crucial role of paradigm functions is to derive each inflected form of a lexeme as a cell in that lexeme's paradigm. So, given the template in (19) and the rule blocks in (25), I define the PF for EP verbal morphology as follows:

(26) PF for EP verbal paradigm

Where X = stem of *L* and σ (= verbal features and clitic features),

PF (<X, σ>) $=_{def}$ RR_I (RR_{II} (X, σ>)) $=_{def}$ <Y, σ>

[44] In Stump's original notation, the Roman numerals are not subscripted. The subscripts given here do not imply any kind of order of application, they only serve to refer to each RR individually. As in the original theory, I assume that the rules apply according to Panini's Determinism (or the principle of the Elsewhere Condition, cf. Chapter 2).

The PF in (26) applies to the pairing <X, σ> (where X is the stem of a lexeme and σ is the complete set of morphosyntactic properties realised by a given EP verb form) to yield the cell <Y, σ> in the lexeme's paradigm. In addition, PFs are crucially defined in terms of RRs. The RRs contained in the rule blocks in (25) are represented as RR_I and RR_{II} where block I applies before block II. Which rules apply from each rule block is determined by the set of features realised by the verb form, as shall be seen below.

Example 1:

Let us now illustrate how ordinary verb forms in EP can be derived within PFM. Given the PF in (26) and the RRs in (25), let us consider the derivation of the 1st person plural Imperfect form of the lexeme LEVAR in (27).

(27) Derivation of *levávamos* ' we took'

Where σ ={Per:1, Num: Pl, Tense: Imperfect}

PF (<leva-, σ>)

a. $=_{def}$ RR_{II} (RR_I (<leva-, σ>)

b. $=_{def}$ RR_{IIb} (RR_{Ia} (<leva-, σ>)

c. $=_{def}$ <leva-va-mos, σ>

The first line specifies the morphosyntactic feature set of the inflected word form; the second line identifies *leva-* as the stem which serves as the morphological base of the suffixes. The Rule Blocks defining the PF are given in (27a) and the RRs from each one of the rule blocks are given in (27b). In the case of *levávamos*, the feature set σ triggers the application of RR_{Ia} from block I and of RR_{IIb} from block II. Based on the rule block index specified on each RR, we assume that RR_{Ia} is realised first and that RR_{IIb} is realised second. We therefore start the derivation by applying RR_{Ia} to the stem *leva-* of LEVAR 'carry' yielding the underived form *leváva-;* and then apply RR_{IIb} to *levá-va-* yielding *levá-va-mos*. An informal representation of the sequential application of realisation

rules is provided in (28). The input to each RR is given in the left column and the output of applying a given RR is shown in the right column.

(28) Cyclic derivation of *levávamos* 'we took'

Input	Output	
a. (leva-)	leva+ *va*	(by RR_{Ia}, 23ai)
b. (levava-)	levava+*mos*	(by $RR_{IIb,}$ 23bii)

Before proceeding with our analysis of EP verb forms, it is useful to clarify why RR_{Ia} is selected and not RR_{Ib} in (27a) and (28a). Note that both RRs derive the imperfect tense marker, but only one is selected. That is because RR_{Ib} in (25aii) provides a more specific realisation, namely the imperfect marker for 2nd Plural verb forms, which in the given derivation cannot apply. So, even though both rules seem to compete for the same affixal slot, only the more general rule can apply (cf. Chapter 2 on Panini's Determinism).

One word is in order about verbal stems. As mentioned above (6.2.1), this thesis assumes that stems are derived through morphomic rules and not through ordinary realisation rules. Therefore a slight change has been introduced into the standard theory of PFM. In Stump (2001), paradigm functions are defined over the root of a lexeme and stem selection is derived through RRs. In particular, the RRs belonging to the first Rule Block are responsible for selecting the stem that is appropriate for a given verb form (Stump 2001, p.45-46). However, following Aronoff (1994) I shall assume that languages have a rich inventory of stem formation rules for the derivation of stems. An example of how such rules may operate is intuitively illustrated in (21) for $stem_1$ and $stem_2$. I shall leave a more detailed treatment of this topic for further research (cf. Appendix A), however I assume that the insight can be accommodated within PFM by defining PFs over indexed stems rather than over indexed roots (Spencer ms, Luís&Spencer in press). The index, of course, ensures that that identity of the lexeme is maintained throughout the derivation. Therefore, the PFs

in (26) and in (27) have been formulated over the stem *leva-* of LEVAR. Crucially, the idea that RRs should apply to 'morphomic' stems rather than to roots of lexemes does not affect the realisational-inferential sprit of the theory.

Example 2

Given the relevance of Future and Conditional verb forms for the analysis of mesoclisis, I shall briefly illustrate how the PFM-analysis applies in the derivation of 1st person Plural forms in both tenses.

Starting with the 1st pl Future form *levaremos*, let us assume that the verb form *levaremos* realises the set of morphosyntactic features given in the first line of (29). Let us also assume that the PF is defined over the stem *levar-* (cf. 21b) and evaluated as in (29a-b) yielding the inflected verb in (29c).

(29) Derivation of *levaremos* 'we will take'

Where σ = {TENSE: Fut; AGR:1PL},
PF (<levar-, σ>)
a. $=_{def}$ RR_{II} (RR_{I} (<levar-, σ>)
b. $=_{def}$ RR_{IIb} (RR_{Ia} (levar-, σ>)
c. $=_{def}$ <levar-e-mos, σ>

The numeral subscript on each rule identifies the order in which they apply. So, RR_{Id}, which supplies the future exponent *-e* marker, applies before the rule RR_{IIb} realising the exponent for 1st pl agreement *-mos*. A word form such as *levaremos* ('we will take') is therefore derived through the layered application of RRs, as illustrated informally in (30).

(30) Cyclic derivation of *levaremos* 'we will carry'

input	output
a. (levar-)	levar+*e* (by RR $_{Id,}$ 23aiv)
b. (levare-)	levare+*mos* (by RR $_{IIb}$, 23bii)

Assuming the RRs in (25), the sequential ordering of affixes in (30) can be described as follows: RR $_{Id}$, in (30a) applies to the infinitival stem *levar*- yielding *levare*-; and RR $_{IIb}$ which derives the suffix –*mos*, in (30b), applies to the output of the previous rule deriving the verb form *levaremos*.

In the derivation of a given verb form, only the narrowest applicable rule applies. So, even though there are two RRs deriving the future tense marker, only the more specific applies. In this case, the RR deriving the future marker -*e* is more specific because it applies to 1st Sg/Pl and 2nd Pl verb forms. Therefore, even though rule (23aiii) would also be applicable at stage (29a), it is pre-empted by the more specific rule in (23aiv) (cf. Chapter 2 on Panini's Determinism).

EXAMPLE 3

Having briefly shownb what the derivation of a Future verb form might look like, I will now consider briefly the analysis of a Conditional form. I shall assume that conditional forms also have two affix-positions: one position for the tense marker and another one for the agreement ending (cf. Appendix A). The stem to which these inflectional suffixes attach is the future/conditional stem introduced in (21b).

Given the affix positions suggested in (21), I propose the RRs in (23e) and (23h) for the derivation of the verb form *levaríamos* 'we would take':

(29) Where σ = {TENSE: Cond; AGR:1PL},
PF (<levar-, σ>)
a. $=_{def}$ RR_{II} (RR_{I} (<levar-, σ>)
b. $=_{def}$ RR_{IIb} (RR_{Ie} (levar-, σ>)
c. $=_{def}$ <levar-ia-mos, σ>

As in the previous derivations, we start by defining the feature set which triggers the set of RRs deriving the inflected word form. We also assume that conditional verbs (as well as future forms) are defined over the infinitival stem of the lexeme. So, in the second line, the PF takes as input the stem *levar*-; (29a) then specifies the RRs which realise the verb's features and which yield the verb form in (29b). The derivation proceeds as shown in (30): we apply RR_{Ie} to the stem *levar*- and derive *levaria*-; we then apply RR_{IIb} to *levaria*- and derive *levariamos*. An informal diagram is given in (30).

(30) Cyclic derivaton of *levariamos* 'we would take'

input	output
a. (levar-)	levar+*ia* (by $RR_{Ic,}$ 23a-iii)
b. (levaria-)	levaria+*mos* (by RR_{IId}, 23b-ii)

Summary:

The aim of section 6.2 was to introduce a fragment of EP verbal conjugation and to show how inflected verb forms are derived within the theory of Paradigm-Function Morphology developed by Stump (1993, 2001). This overview shall facilitate the discussion about cliticisation provided next.

6.3 Analysis: first proposal

Let us now turn to the derivation of enclisis and mesoclisis. As I have done for tense and agreement suffixes, I will start by examining the realization of clitic sequences within the theory of Paradigm-Function Morphology (Stump 2001), using insights from Spencer (2000). In 6.3.1, the cluster template shall be examined and RRs shall be provided. This is followed by a preliminary PF-analysis of cliticised verb forms in 6.3.2 and by a revision of that analysis in 6.3.3. Section 6.3.4 finally offers a brief summary of previous approaches to enclisis and mesoclisis.

6.3.1 Rule blocks

Before determining the exact number of realisation rules that are necessary for the derivation of suffixing clitics, we will need to examine how many clitic positions there are in EP. Note that for each position, we will assume one rule block which provides all the affixes that fill that position.

As mentioned in Chapter 4, the clitic system of EP comprises reflexive, dative and accusative clitic pronouns which combine with each other in rigid order: reflexive clitics precede dative clitics and dative clitics precede accusative clitics. Rigid ordering is one of the properties of clitic sequences which most clearly reflects similarities with inflectional affixes. In inflectional morphology, affix linearisation is obtained by organising RRs into rule blocks (cf. ch2). Given the principle that affix order reflects the order in which rule blocks apply, one might propose the following rule blocks for EP clitics:

(31) Block $_A$ (reflexive clitics): me, te, se , nos, vos, se
Block $_B$ (dative clitics): me, te, lhe, nos, vos, lhes
Block $_C$ (accusative clitics): me, te, o/a, nos, vos, os/as

By (31), then, block A supplies reflexive clitics, block B supplies dative clitics and block C supplies accusative clitics. Since rule block A applies before rule block B, and block B before block C, we predict that reflexives must always precede datives and that datives must appear before accusatives, thus capturing clitic linearisation without additional stipulations.

The problem with the rule blocks in (31) however is that each block realises all person and number combinations for the case features it is associated with: we have 1st, 2nd and 3rd person singular and plural clitics in each one of the rule blocks. This is a problem because rule block ordering is based on the assumption that affixes from different rule blocks may co-occur within the same word form (unless of course co-occurrence restrictions are formulated which explicitly rule out

certain combinations). Given (31) then, one would predict that each clitic suffix from block B may co-occur with each clitic suffix forms block C, for example. This, of course, generates far more clitic combinations than are actually attested. Recall that one of the co-occurrence restrictions found in EP clitic combinations prevents 1st person and 2nd person clitic (regardless of case) from co-occurring (cf. ch.4). Thus combinations such as *me-me, *me-te, *te-me, *nos-vos, *me-vos, etc. are not attested, regardless of whether they are Refl-Dat or Dat-Acc clusters.

The number of clitic combinations predicted by (31) can be significantly reduced by rearranging the rule blocks, thereby preventing certain clitics from co-occurring. Based on the assumption that affixes belonging to the same block cannot co-occur (because each rule block can only apply once in the derivation of a given word form, cf. Anderson 1992, Stump 2001), we can assign all 1st and 2nd person clitics to the same rule block, thus predicting that 1st and 2nd person clitics will never combine with each other. One provisional proposal is shown below where I assume four affixal slots for clitics:

(32) Tentative proposal for clitic rule blocks

A	B	C	D
Ref 3	1&2	Dat3	Acc3

(32) assigns 1st and 2nd person clitics to block B (regardless of case). This effectively means that a) block A would realise only reflexive 3rd person clitics (regardless of number), b) block C would supply rules for 3rd person dative clitics; and that c) block D would be restricted to 3rd person accusative clitics.

There is however one further restriction on clitic combinations which (32) leaves unaccounted for and which can be easily captured through rule block ordering: namely the fact that only 3rd person reflexives can combine with dative clitics (e.g, se-me, se-lhe, but not *me-lhe or *te-lhe). This regularity can be nicely derived if we reduce block B and C to just one rule block, if we assume that

affixes belonging to the same rule block cannot co-occur. Based on Luís (2001a), I shall propose the rule blocks in (33):

(33) Clitic rule blocks for EP

A	B	C
Ref 3	1&2, Dat3	Acc3

The template in (33) appears to predict that a) only Ref3 clitics can combine with dative clitics, b) dative clitic can only combine with accusative 3rd person clitics, and c) 1st and 2nd clitics cannot co-occur[45].

6.3.2 Clitic suffixation

Having discussed the number of rule blocks and the order in which they apply, I will now address the realisation rules provided by these rule blocks. I propose that realisation rules should be employed for the derivation of clitic affixes, however I also assume these RRs take as input an underived base rather than a fully formed verb. Let us recapitulate somewhat the motivation underlying this proposal.

As alluded to before (cf. 6.1), accounting for enclisis in EP is complicated by the fact that clitic suffixes appear verb-internally. Enclisis could be simply regarded as affixation to an inflected verb form as standardly assumed for various languages, however, the phenomenon of mesoclisis requires a more subtle derivation of enclisis because we must explain why one and the same clitic suffix in EP can appear before ordinary tense and agreement suffixes.

[45] Of course, not all restrictions found in inflectional systems can be captured through rule block ordering (cf. Stump 2001, Chapter 2). The fact that position I cannot combine with position III will be reinforced through co-occurrence restrictions.

As a first approximation, I propose the PF in (34) for the derivation of the affix sequence of verb forms with verb-final clitics as shown in (35).

(34) General PF for cliticised verb forms in EP (provisional)

Where X = stem of *L* and σ (= verbal features and clitic features),

PF (<X, σ>) $=_{def}$ RR_C (RR_B (RR_A (RR_{II} (R_I (<X, σ>))))) $=_{def}$ <Y', σ>

(35) a. leva-vá-mos-lhe ('we took for him/her')
stem-Impf-1Pl-dat.3Sg

b. apresenta-va-m-se-lhe ('they introduced themselves to him')
stem-Impf-3Pl-refl.3Pl-dat.3Sg

This PF is defined both in terms of ordinary realisation rules - namely RR_I, RR_{II} - which derive the exponents for tense and agreement (cf. 6.2) - and in terms of the realisation rules which derive clitic suffixes - namely RR_A, RR_B, RR_C - (cf. 6.3.1). It is further assumed that rule blocks I and II realising tense and agreement suffixes precede all clitic rules, thus capturing the fact that clitic suffixes appear after all ordinary suffixes. We shall later see why it is important to jointly specify both sets of rules. As will be shown below, this procedure will enable us to realise clitic suffixes in reversed order.

In what follows, I formulate the realisation rules deriving clitic suffixes (6.3.2.1) and provide a preliminary analysis of enclisis and mesoclisis based on a modest extension of the format of Paradigm-Functions (6.3.2.2).

6.3.2.1 Realisation rules

As alluded to before, RRs within PFM derive affixes as the phonological output of a given set of morphosyntactic features (cf. ch.2). The complete feature set expressed by EP pronominal clitics is summarised below:

(36) Clitic features and permissible values

CASE: Accusative, Dative
REFLEXIVITY: yes, no
PERSON: 1,2,3
NUMBER: Singular, Plural
GENDER: Masculine, Feminine

Based on these features and assuming the rule block order suggested in (37), I propose the following set of the rules for the generation of clitic suffixes:

(37) Clitic rule blocks (provisional)

a. Block A

i. $RR_{A1,\ \{Person:3,\ Refl:+\},\ V}$ (<X,σ>) $=_{def}$ <X-s*e*, σ>

b. Block B

i. $RR_{B1,\ \{Person:1,\ Number:\ Sg\},\ V}$ (<X,σ>) $=_{def}$ <X-*me*, σ>

ii. $RR_{B2,\ \{Person:2,\ Number:\ Sg\},\ V}$ (<X,σ>) $=_{def}$ <X-*te*, σ>

iii. $RR_{B3,\ \{Person:1,\ Number:\ Pl\},\ V}$ (<X,σ>) $=_{def}$ <X-*nos*, σ>

iv. $RR_{B4,\ \{Person:2,\ Number:\ Pl\},\ V}$ (<X,σ>) $=_{def}$ <X-*vos*, σ>

v. $RR_{B5,\ \{Case:\ Dat,\ Person:3,\ Number:\ Sg\},\ V}$ (<X,σ>) $=_{def}$ <X-*lhe*, σ>

vi. $RR_{B6,\ \{Case:\ Dat,\ Person:3,\ Number:\ Pl\},\ V}$ (<X,σ>) $=_{def}$ <X-*lhes*, σ>

c. Block C

i. $RR_{C1, \{Case: Acc, Person:3, Number: Sg, Gender: Masc\}, V}$ (<X,σ>) $=_{def}$ <X-*o*, σ>

ii. $RR_{C2, \{Case: Acc, Person:3, Number: Sg, Gender: Fem\}, V}$ (<X,σ>) $=_{def}$ <X-*a*, σ>

iii. $RR_{C3, \{Case: Acc, Person:3, Number: Pl, Gender: Masc\}, V}$ (<X,σ>) $=_{def}$ <X-*os*, σ>

iv. $RR_{C4, \{Case: Acc, Person:3, Number: Pl, Gender: Fem\}, V}$ (<X,σ>) $=_{def}$ <X-*as*, σ>

Rule (37ai) realises the 3rd person reflexive clitics, both plural and singular, and is intended to fill the first clitic position; to capture the syncretism between singular and plural forms I assume that there is only one RR realising reflexive 3rd person clitics which is unspecified for number.

In (37b), the RRs realise 1st and 2nd person clitics and 3rd person dative clitics, and fill in the second clitic position in (33). Because of the case syncretism displayed by 1st and 2nd person clitics, I leave the RRs (37bi-iv) unspecified for case and assume that only one RR is needed to generate each person and number combination for reflexive, dative and accusative forms (a more detailed discussion about clitic inventory is given in Chapter 9). Finally, the rules in (37c) realise 3rd person accusative clitics for all number and gender combinations and fill the third affix position in (33).

6.3.2.2 Preliminary proposal

Based on the assumption that clitic suffixes are derived though the set of RRs presented in (37), this section now examines the derivation of enclisis and mesoclisis.

a) Enclisis

Let us now illustrate how the PF in (34) would derive the cliticised verb form *levávamos-lhe*. The morphosyntactic properties of this verb form are given in the first line in (38); they include the ordinary features and the clitic features. In the second line, the PF is defined over the stem *leva-*, as suggested for the derivation of the ordinary verb form *levávamos* in (27). This PF is evaluated in (38a) by the RRs in (37) yielding the inflected verb in (38b).

(38) Where σ = {Tns: Impf, Agr:2Pl} and {Case:Dat,Per:3,Num:Sg},

PF (<leva-, σ>)

a. $=_{def}$ $RR_{C\{\}}$ (RR_{B5} ($RR_{A\{\}}$ (RR_{IIb} (R_{Ib} (<leva-, σ>)))

b. $=_{def}$ <leva-va-mos-lhe, σ>

We start the derivation of *levavamos-lhe* by applying RR_{Ib} to the infinitival stem *leva-* of LEVAR, yielding the underived form *levava-*. The subsequent rule, namely RR_{IIb}, applies to the output of the previous rule, realising the agreement ending *–mos* and yielding the form *levavamos-*. Note that because the PF defines the application of further rule blocks for the derivation of clitic suffixes, this verb form does not constitute a fully derived form yet. It shall instead serves as the morphological base for the suffixation of clitics. Let us then see how the RRs from the rule blocks A, B And C are evaluated.

Given that that only a maximum of two clitics can ever occur in cliticised verb form in EP, we predict that cliticised verbs shall contain one or two empty slots. Empty slots of this type results when the set of morphosyntactic features of a given cliticised verb requires specific slots to be realised without any exponents. If for example a given verb is specified for 3rd dative features, such as in (38), we predict that neither rule block A nor C will provide a rule because they do not contain any rule realising a subset of σ of the word's features. In this case then the rule block will not supply an exponent and leave the slot 'empty'. This is what happens with the cliticised verb in (38): no RR from block A is applicable because none of the rules in block A (cf. 37) realises a subset of σ (in 38). By the Identity Function Default (cf. ch.2), then, this rule block supplies no exponent and must be evaluated as the identity function. As mentioned before, this is a function which, when applied to an argument, delivers that argument as its value: f(X) = X. In other words, the input to $RR_{A\{\}}$ in (38a) is also its output, namely *levavamos-*:

(39) By the Identity Function Default, $RR_{A\{\}}$ (<levavamos-, σ>) $=_{def}$ <levavamos-, σ>

Rule block B, on the contrary, provides RR_{B5} realising the features {Dative, 3Sg} which constitute a subset of the features given in (38). This rule applies to *levavamos-* yielding the cliticised verb *levavamos-lhe*. Finally, rule block C (like rule block A) shall be realised vacuously because, once more, no explicit rule from this block can apply in this derivation. By the Identity Function Default, then, no exponent is supplied. A diagrammatic representation of the derivation is given below:

(40) Derivation of *levavamos-lhe*

input	output
a. leva-	leva+*va* ($RR_{Ia,\ by}$ 23a-i)
b. levava-	levava+*mos* ($RR_{IIb,\ by}$ 23b-ii)
c. levavamos-	levavamos+{} ($RR_{A,}$ vacuous)
d. levavamos-{}	levavamos+{}+*lhe* (by $RR_{B5,}$ 34b-v)
e. levavamos-{}-*lhe*	levávamos+{}+*lhe*+{} ($RR_{C,}$ vacuous)
	levávamos-lhe

The sketch in (40) constitutes an attempt at capturing the linearisation of suffixing-clitics and their status as verbal affixes. Cliticised verb forms are derived through the piecewise application of two sets of RRs, as the 'input' and 'output' forms of the derivation of *levávamos-lhe* above suggest.

b) Mesoclisis

Assuming that the sequence given in (40) presents the default affixal sequence, let us now consider mesoclisis. Let us assume for the moment that mesoclisis follows from a restriction on the general PF given in (41) which re-arranges the order in which clitic suffixes are realised placing them before the tense and agreement marker in future and conditional verb forms. This type of rule block reordering might be represented as in (41). We are here essentially adopting Stump's proposal for Fula 'reversed position classes', summarised in 6.1.3.

(41) Statement for reversed clitic order for EP (provisional)

Where σ = {Tns: Fut/Cond}and {Case: α }

PF (<X, σ>)

a. = $_{def}$ $\mathbf{RR_C}$ ($\mathbf{RR_B}$ ($\mathbf{RR_A}$ (RR_{II} (R_I (<X, σ>)))))

b. = $_{def}$ RR_{II} (R_I ($\mathbf{RR_C}$ ($\mathbf{RR_B}$ ($\mathbf{RR_A}$ (<X, σ>)))))

The first line of the statement specifies the morphosyntactic context within which rule block reversal takes place (cf. ch. 9 for refinement); the second line introduces the PF and specifies the expression which serves as its input; the third line defines the default rule block order which realises clitic suffixes after tense and agreement suffixes; finally, the last line defines the 'reversed' rule block order placing clitic suffixes before all ordinary suffixes. To understand the full extent of this provisional proposal it is important to recall that PFs determine the order in which affixes are realised. For EP, the default order places ordinary affixes before clitics; for mesoclisis, this order would be reversed. A sample derivation is given below for *levar-lhe-emos* ('we will take for him/her'). The feature set realised by this verb form is {Tns: Fut, Per:2, Num:Pl} and {Case: dat; Person: 3, Number: sg}. As shown below:

(42) Where σ = {Tns: Fut, Per:2, Num:Pl} and {Case: dat; Person: 3, Number: sg}),

PF (<levar, σ>)

a. = $_{def}$ $\mathbf{RR_{C\{\ \}}}$ ($\mathbf{RR_{B5}}$ ($\mathbf{RR_{A\{\ \}}}$ (RR_{II} (R_I (<levar, σ>))))) (by 36)

b. = $_{def}$ RR_{IIb} (R_{Id} ($\mathbf{RR_{C\{\ \}}}$($\mathbf{RR_{B5}}$ ($\mathbf{RR_{A\{\ \}}}$ (<levar, σ>))))) (by 46)

c. = $_{def}$ <levar-lhe-e-mos, σ>

(42a) defines the PF over the infinitival stem of the verbal lexeme, in this particular case *levar-* of the lexeme LEVAR. The default order of rule block application is given in (42b) following the PF provided in (34). (42c) then defines the reversed order for rule block application and determines that

all clitics are realised before the tense and agreement markers. An informal illustration of the cyclic application of RRs is given below:

(43) *levar-lhe-emos* 'we will carry for him/her' (informal)

input	output	
a. levar-	levar+{}	(RR_A, vacuous)
b. levar-{}-	levar+{}+lhe	(RR_{B5}, by 34b-v)
c. levar-{}-lhe-	levar+{}+lhe+{}	(RR_C, vacuous)
d. levar-{}-lhe-{}-	levar+{}+lhe+{}+e	($RR_{Id,}$ by 23a-iv)
e. levar-{}-lhe -{}-e -	levar+{}+lhe+{}+e+mos	($RR_{IIb,}$ by 23b-ii)
	levar-lhe-emos	

Comparing the derivation of mesoclisis in (43) with that of enclisis in (40), it becomes evident that there is a difference at the level of ordered rule block application. What we see is that for the derivation of mesoclisis one would need to realise clitics before the tense and agreement markers. This entails that clitics attach to the infinitival stem of the verbal lexeme, in this particular case it is the stem *levar-* of the lexeme LEVAR (43a), to which the tense and agreement normally attach in non-cliticised future/conditional verb forms (cf. 6.2).

Summing up, the above discussion has shown that the joint specification of both sets of affixes assign suffixing clitics the ability to appear verb-internally by means of rule block reordering.

6.3.2.3 Discussion

Even though the analysis in (42) correctly derives mesoclisis as a case of rule block reordering, it fails to provide a principled explanation for the fact that verb-internal clitics appear before 'all' ordinary suffixes and not just before some. So, verb forms in which clitics appear between the tense marker and the agreement marker are not attested, but the 'reordering' suggested above does not explain why there is such a restriction:

(44) levar -lhe -e -mos → *levar-e-lhe-mos
take -dat.3sg -fut -1pl

If reordering applies to individual rule blocks, then it should be technically possible to derive the affix order in (44). However, only two positions are available: either verb-final or immediately adjacent to the infinitival stem. What we need, of course, is to capture the insight that ordinary suffixes cannot be broken up[46].

Directly related to this issue is also the next problem. By lining up one rule block after another, the PFs in (34) and (41) fail to explain why a clitic moves as a whole unit and why it is completely impossible in EP (and most Romance languages) to leave one of the clitic suffixes behind[47]. So, what we need is to explain why 'affix-reordering' affects the complete set of clitic suffixes and not just part of it, as shown below:

(45) levar -no -lo -ia → *levar-nos-ia-lo
take -dat.2pl - acc.3sg.masc -cond.1/3sg

Of course, it might be possible to stipulate that reordering affects the whole sequence of clitics, that however would still not explain why the cluster appears before the tense and agreement marker and not between each one of these ordinary suffixes as in (44).

What examples (44) and (45) effectively show is quite important, namely that the interaction between sequences of affixal clitics and sequences of ordinary affixes is quite restricted in EP: as in most of the other Romance languages, cannot interrupt sequences of ordinary affixes (44) and, inversely, ordinary affixes cannot intervene between clitic clusters (45). This effectively indicates

[46] Note that the idea of defining future and conditional forms over an infinitival stem automatically prevents clitics from appearing between the theme vowel and the '-r' suffix. The stem is derived at the morphemic level, through morphemic rules, not through ordinary realisation rules.

[47] An exception to this behaviour are split clitic clusters in Standard Modern Greek (Terzi 1999).

that sequences of affixes and sequences of clitics do not interleave; they are therefore best viewed as distinct (though co-occurring) layers of inflectional morphology (further evidence supporting this view is given in Chapter 9 where I shall discuss clitic ordering and preverbal placement in more detail). We shall return to clusters in Chapter 9 (cf. also Anderson 1995, Monachesi 1999, Gerlach 2001a).

Even though the analysis of affix-metathesis has been inspired by Stump's analysis of Fula (6.1.2), it has now become clear that there is one significant difference between the Fula data and EP mesoclisis. It is the fact that Fula reordering affects two adjacent affixes while EP mesoclisis affects whole sequences of affixes - tense and agreement suffixes, on the one hand, and clitic suffixes, on the other. By simply extending Stump's analysis to EP we are overlooking the fact that the degree of locality which characterises the Fula data is absent from EP; in fact, mesoclisis cannot just be analysed as a typical case of affix reversal. Nonetheless, the intuition that EP affixes have the ability to 'change places' seems to be essentially right. The question now is how to capture the interaction between ordinary suffixes and clitic suffixes attested for EP.

6.3.3 Revised analysis

We have seen that the format of the PF deriving cliticised verbs in (34) cannot capture the intuition that both sequences of clitics and sequences of ordinary suffixes behave like unified wholes. The fact that this generalisation is completely missed suggests that an alternative analysis is needed. It is however not entirely obvious how to best capture the limited interaction between ordinary affixes and affixal clitics. So far, these issues have not been addressed in much detail because it is generally assumed that clitics attach to fully formed verbs. The analysis therefore shall be based on a a modest extension of the standard formulation of PFs. In the first part a revised format of PF will be suggested and illustrated with a case of verb-final cliticisation (6.3.3.1). The second part then addresses mesoclisis and offers an analysis based on the revised format for PFs (6.3.3.2).

6.3.3.1 Verb-final suffixation

For the analysis, I shall propose the revised PF given in (46). This is a PF defining two sets of rule blocks, each realising its own set of morphosyntactic features.

(46) PF for cliticised verbs

Where X = stem of *L* and $\sigma = \sigma_1$ (= verbal features) and σ_2 (= clitic features),

PF (<X, σ_{1+2}>) $=_{def}$ $RR_A (RR_B (RR_C(\sigma_2))) \circ RR_{II} (RR_I (\sigma_1))$ (X)

What we have in (46) is therefore an 'extended' PF which derives cliticised verbs as the output of two subsets of features, namely σ_1 for ordinary verbal features and σ_2 for clitic suffixes. The important point about the paradigm function is that each set of RRs realises its own set of features, taking as input one and the same morphological base X. Given that there are two sequences of rules, the PF also captures the fact that each set of rules forms its own layer of affixation.

An alternative way of visualising the PF in (46) would be to adopt a more 'vertical' representation of the PF as in (47) which shows that we are dealing with embedded layers of affixation.

(47) PF (<X, σ_{1+2}>) $=_{def}$

base	i. stem:	X
	ii.verbal suffixes:	$RRI(RRII(\sigma_1)$
clitic cluster:		$RR_A (RR_B (RR_C(\sigma_2)$

In this 'vertical' format of the PF, the inner layer corresponds to the 'clitic base'; the outer layer to the sequence of clitic suffixes. For reasons of clarity and exposition, I shall illustrate the derivation using the format in (47).

Starting with the sequence of ordinary RRs, namely RR_{II} and RR_I, these apply cyclically to a basic (underived) stem X. By underived (or basic) stem I understand the stem which contains a root

and a theme vowel (cf. 6.2). The combination of this 'basic' stem with the suffixes yields the inflected verb form which serves as input to the RRs deriving the clitic suffixes. This first stage of the derivation would be roughly equivalent to the derivation of ordinary verb forms and could be easily represented as an inner PF which is embedded inside the outer layer of affixation which realises the feature set σ_2 and is realised through the rule blocks $RR_{A,}$ RR_B, and RR_C [48].

A sample derivation of enclisis is given in (48). The features associated with the cliticised verb are given in the first two lines: σ_1 for ordinary affixes and as σ_2 for pronominal features. The stem over which the PF is defined is *apresenta-* of the lexeme APRESENTAR. Let us then evaluate the PF in a stepwise fashion: (48a) defines two sets of rules, one deriving ordinary verbal suffixes and another deriving clitic suffixes. Starting with the ordinary rules, the cyclic application of the ordinary RRs yields the verb form *apresentavam* in (48b) which serves as the clitic base of (or input to) the clitic string. From the application of the RRs deriving the clitic cluster, we derive the complete cliticised verb form *apresentavam-se-lhe* in (48c). As mentioned before, if a given rule block does not realise any subset of features, then the Identity Function Default applies (6.2.2). This is what happens with rule block C.

[48] Here it is assumed that realization rules deriving the clitic sequence apply successively. This view however shall be revised in Chapter 9.

(48) Derivation of *apresentavam-se-lhe* 'they introduced themselves to him/her'

Where $\sigma_{1 =}$ {Tns:Impf, Per:3, Num:PL;} and $\sigma_{2 =}${[Case:Ref,Per:3, Num:Pl]; [Case:Dat,Per:3,Num:Sg]},

PF (<apresenta-, σ_{1+2}>) $=_{def}$

a. $=_{def}$ { base: { i. stem$_1$ of L: apresenta- ; ii.verbal suffixes: RR_{IId} (RR_{Ia} (σ_1) } ; clitic cluster: $RR_{C\{\}}$ (RR_{B5} (RR_{A1} (σ_2) }

b. $=_{def}$ { base: apresenta-va-m ; clitic cluster: $RR_{C\{\}}$ (RR_{B5} (RR_{A1} (σ_2) }

c. $=_{def}$ <apresentavam-se-lhe, σ_{1+2}>

6.3.3.2 Verb-internal suffixation

Having addressed the default placement of clitic suffixes, we will now examine the derivation of verb-internal cliticisation. The affix order we need to derive is given in (49).

(49) *levar-lhe-e-mos* '(we) will take for him/her'

i.	*levar*	*-lhe*	*-e*	*-mos*
ii.	stem	Dat.3.sg	Future	Agr
iii.	X	**RR_{B5}**	RR_{Id}	RR_{IIb}

We have seen in 6.3.2 that by reordering clitics on a one-by-one basis it is very difficult to explain why clitic sequences never break up the sequence of ordinary affixes. This suggested to us that an account of mesoclisis should be able to capture two aspects about the phenomenon: first, the fact that enclitics such as *-lhe* or *-se* can appear verb-internally before ordinary affixes; second, the fact that clitics never occur between the tense and agreement marker.

At this point let us recall the basics of the analysis of Fula where the main claim was that PF, which determine the order in which RR apply, can be subject to rules changing the default order (cf. 6.1.2). We shall now extend this idea to the PF in (47) which, as we alluded to before, is defined in terms of two layers of affixation. We therefore propose the rule in (50) in which mesoclisis results from reversing the order in which each layer applies.

(50) Reversed PF for verb cells with verb-internal clitic suffixes

Where σ_1 = {Tns: Fut/Cond} and σ_2{Case: α; Person: β, Number: γ},

PF (<X, σ_{1+2}>) $=_{def}$ RR_{II} (RR_I (σ_1)) ◦ RR_A (RR_B (RR_C(σ_2))) (X)

This insight can also be visualized diagrammatically in (51):

(51) Where σ_1 = {Tns: Fut/Cond} and σ_2{Case: α; Person: β, Number: γ},

PF (<X, σ_{1+2}>)

$=_{def}$

base:	i. $stem_2$ of L:	X-
	ii. clitic cluster:	RR_C (RR_B (RR_A (σ_2)
verbal suffixes:		RR_{II} (RR_I (σ_1)

(51) shows that - if a verb's σ_1 set of features is specified for Future or Conditional tense - the sequence of ordinary suffixes is realised after the sequence of clitic suffixes. In more concrete terms, this amounts to saying that the rule blocks realising clitics apply before the rule blocks realising the ordinary verbal affixes.

Let us now see how the PF is evaluated for the verb form *levar-lhe-emos* ('we.pl will take for him/her'):

(52) Where σ_1 ={Tns:Fut,Per:1,Num:PL} and σ_2 ={[Case:Dat,Per:3, Num:Pl]

PF (<levar, σ_{1+2}>)

a. $=_{def}$ { base: { i. $stem_2$ of L: levar- ; ii. clitic cluster: $RR_{C\{\}}$ (RR_{B5} ($RR_{A\{\}}$ (σ_2) } ; verbal suffixes: RR_{IIb} (RR_{Id} (σ_1) }

b. $=_{def}$ { base: levar-lhe ; verbal suffixes: RR_{IIb} (RR_{Id} (σ_1) }

c. $=_{def}$ <levar-lhe-e-mos, σ_{1+2}>

The evaluation proceeds as follows: (52a) defines the PF in terms of its stem and in terms of the rules realizing both ordinary and clitic suffixes. As to the set of clitic-rules, in (52a), only block B contains a rule that can realise a subset of the verb's complete feature set. Block A and C apply vacuously. Starting with the derivation of the clitic suffix, the application of the clitic rule RR_{B5} to the stem *levar-* yields the cliticised stem *levar-lhe* in (52b). Through the successive application of the verbal suffixes to the cliticised stem, we obtain the cliticised verb form *levar-lhe-emos* in (52c).

What the analysis shows is that by reversing the order of rule blocks we are effectively saying that mesoclisis, like enclisis, is suffixation to a verbal stem.

Summary

This section formulated a Paradigm Function for the derivation for verbs with suffixing clitics. It has been argued that the PF for cliticised verbs must define both the feature set for ordinary verbal exponents and the set for clitic exponents. This claim has been motivated with data from EP mesoclisis which showed that verb-internal cliticisation must be analysed as affix reordering. The

need to be explicit about both verbal features and clitic features is therefore an important claim which previous inflectional studies of clitic phenomena have not yet proposed.

This proposal constitutes a slight extension of the original theory of PFM, but it endorses the realisational spirit of the theory of PFM and preserves the crucial role of Paradigm Functions as functions which determine the inflectional paradigm of a given lexeme. Given this analysis, PFs constitute functions which determine the complete set of cliticised forms of a verbal lexeme through the application of two sets of Realisation Rules (RRs). The analysis therefore differs from that of Spencer (2000) for Macedonian clitics in that RRs deriving clitics are not derived over fully formed verbs (53).

(53) Format for extended paradigm function (EPF)

$$\text{EPF}(\langle V, \sigma\rangle) = _{\text{def}}\ RR_{IV}(RR_{III}(RR_{II}(RR_{I}(\langle V, \sigma\rangle))))$$

$$= _{\text{def}}\ (\langle W, \sigma\rangle)$$

The Extended Paradigm Function formulated by Spencer (2000) says that the PF deriving affixal clitics in Macedonian takes as input (<V, σ>), where σ stands only for the set of clitic features associated to an inflected verb and where V is the cell from the verbal paradigm of this language to which clitics attach. For EP, however, the Extended Paradigm Function as formulated in (53) cannot be adopted, because it regards the clitic base as an opaque morphological host whose internal parts (i.e., affixes) cannot be manipulated (i.e., reordered). The analysis sketched here instead 'decomposes' the V unit in (53). In this respect, both accounts capture the insight that cliticisation should be regarded as a type of 'extended inflection' (Zwicky 1986, Halpern 1995) and that cliticised verbs build the 'extended' verbal paradigm of a language.

6.3.4 Previous accounts

The idea that EP postverbal clitics could be viewed as verbal inflections is also found in Zwicky (1987), Stolz (1992), Spencer (1992). These studies however do not formalise their claims and thus no comparison can be made with the present proposal. Other accounts of EP cliticisation are either syntactically or phonologically inspired, although attention is given to the non-productive aspects of enclisis and mesoclisis (Vigário 1999b, Duarte&Matos 2000, Leeuw 1997, Gerlach 2001a). The work of Crysmann (1997, 2002) is an exception, given that it constitutes an attempt at deriving clitics as hybrids, partly syntactic and partly morphological units. In mu survey, I shall restrict myself to Crysmann (1997, 2002) and Vigário (1999b) (cf. Chapter 3 for an general survey; cf. Chapter 7 and 8, for more specific surveys).

6.3.4.1 Crysmann's (2002) 'Hybrid' affixes

As mentioned before (cf. Chapter 3), Crysmann (2002) generates clitic pronouns as syntactically transparent affixes, or 'hybrids', that are available to syntactic ordering. It seems important to point out that, even though clitics are claimed to be affixes, they are assumed to introduce their own domain object, like genuine words. Because of the 'liberated' status enjoyed by clitics, Crysmann's analysis cannot predict that enclitics form a morphologically cohering unit with the verb; this crucial property of EP clitic pronouns must instead be stipulated. As to mesoclisis, Crysmann takes the view that they should be derived as a lexical units. Yet again, the 'synthetic' nature of mesoclitic verb forms only results if we accept Crysmann's claim that separate domain objects can form one lexical unit. The treatment is therefore not typically inflectional and it is at times difficult to convert his theoretical assumptions into a more conventional models of lexicalist grammar.

In the specific case of mesoclisis, it is worth noting that one further technicality is introduced to obtain the correct liner order of affixes. It is argued that tense/agreement endings, such as *-emos* 'fut.2pl' or *-íamos* 'cond.2pl', constitute affixal units attaching to the right-edge of the verb,

regardless of whether the future form is cliticised or not. The analysis then assumes that the reference point for placing (fut&cond) tense/agreement is the outer edge of a given word boundary. The problem however is that there is little motivation for adopting this view within the Portuguese verbal paradigm: first because tense-agreement endings in non-cliticised forms behave in all respects like tense and agreement markers in other verbal tenses; second, because the so-called tense/agreement unit constitutes in effect two concatenated suffixes (with a few exceptions) which appear recurrently throughout the verbal paradigms of EP preceding (not following) enclitics (e.g., vend-ía-mos-lhe 'we used to sell to him/her'). On the contrary, the analysis proposed above, derives affix order in cliticised and non-cliticised verb forms in exactly the same way, namely as suffixes.

From a cross-linguist viewpoint, the analysis proposed by Crysmann also entails the idea that EP clitic pronouns are quite unlike clitics in other Romance languages. Drawing an analogy with German separable-particle verbs and incorporating German 'field theory' into Portuguese syntax, it is argued that a 'new' affixal hybrid is necessary to account for the EP data. Of course the problem with this view is that the numerous similarities EP shares with affixal clitics in French (Miller&Sag 1997, Auger 1994) and Italian (Monachesi 1999), as chapter 4 has argued, turn out to be pure accident. On the contrary, under the inflectional analysis developed in this chapter, we assume that there is a common inflectional basis to Romance cliticisation and that the typological diversity within Romance can be accounted for under the same analysis. This idea will become clearer in chapter 9.

6.3.4.2 Vigário (1999b)'s Phrasal Phonology approach

In Vigário's analysis, as mentioned in Chapters 3 and 5, clitic pronouns are regarded as function words, and the verb-enclitic unit is analysed as a postlexical (i.e., phrasal) combination. Arguments supporting this view are mostly phonological, however the previous chapter has shown that the

evidence in favour of a prosodic analysis is quite weak. Likewise, Vigário's analysis of the phonological idiosyncrasies also raises several problems as shall be discussed in Chapter 7.

The derivation of enclisis has been summarized in Chapter 3. Let us now address Vigário's approach to mesoclisis. Mesoclisis is also assumed to follow from a syntactic derivation. A verb form such as *levar-lhe-emos* is essentially regarded as the combination of three syntactic units: the infinitive verb, the clitic pronoun and the so-called tense/agreement unit. This view, which reflects the historical origin of mesoclisis, raises several problems.

First, if the mesoclitic is assumed to attach syntactically to the verb, then the wide range of morphophonological alternations triggered by it are left unaccounted for. As pointed out earlier in Chapter 4, post-verbal as well as verb-internal clitics induce consonant deletion on the verbal host and undergo phonological change. These shape variations are grammatically conditioned, leaving therefore no doubt about their purely morphophonological nature.

Second, deriving a verb form such as *levar-lhe-emos* as a combination of three syntactic units necessarily entails two doubtful claims: a) that *levar-lhe* is an encliticised infinitive form and b) that *levar-lhe-emos* is the 1st plural future of a cliticised infinitive. Of course, neither one nor the other captures that *levar-lhe-emos* is a cliticised future verb form. As mentioned before, the form *levar* which resembles a non-finite verb form is in fact a non-finite *stem*, used for its form, but not for its grammatical value (cf Aronoff 1995, on parasitic stems). With this function, however, it can only be used as the morphological base of a word, not as a syntactic unit. The mere fact that the base to which the mesoclitic attaches is 'morphomic' (form without content, Aronoff 1995) immediately suggests that the future/conditional & agreement markers are inflectional markers of that infinitival stem and cannot be syntactically separated from the verb. From here it follows that the verb-internal clitic can only be derived as a word-internal suffix.

Additional problems derive from the claim that tense/agreement endings constitute an auxiliary-like unit. It is assumed that future/conditional endings form with agreement endings a syntactic

word-level unit (with unclear properties). This view, of course, reflects the diachronic origin of the future and conditional verb forms, but it is at odds with current synchronic facts. For example, it fails to explain why there is a set of tense and agreement combinations which are exactly identical to the so-called auxiliary unit. Also, the fact that this putative unit only ever occurs verb-finally, without any kind of syntactic mobility, seems to weaken its status as an auxiliary-like unit. Most of the problems just alluded to are generally common to most accounts which do not derive mesoclisis as an inflectional phenomenon (cf. Leeuw 1997, Gerlach 2001a, Duarte et al. 1995). None of these problems however are encountered if postverbal clitics, both enclitics and mesoclitics, are viewed as clitic suffixes.

6.4 Summary

This chapter explored an inflectional analysis of clitic suffixation, both in verb-final and verb-medial position, adopting a slightly extended version of the theory of Paradigm Function Morphology. The position was taken that enclisis should be regarded as verb-final suffixation, while mesoclisis constitutes suffixation to an infinitival stem. Viewing postverbal clitics as verbal morphology captures the data examined in Chapters 4 and 5 about the affixal behaviour of enclitics.

Our analysis of the data showed that clitic suffixes in EP cannot apply to a fully formed verb, as generally assumed in inflectional approaches to cliticisation in Spanish, Italian, French and Macedonian (cf. 6.2.1). The fact that EP suffixal clitics can appear either verb-internally or verb-finally indicates that tense and agreement affixes, on the one hand, and clitic suffixes, on the other, must be able to 'invert' their order of occurrence. Within the theory of Paradigm-Function Morphology, it was argued that the paradigm functions define cliticised verbs in terms of two layers of inflection (cf. Halpern 1995, Spencer 2000): one layer appears for ordinary suffixes and another for clitic suffixes. It was further argued that paradigm functions determine the order in which these

layers can apply. While enclisis attaches clitic suffixes *after* all ordinary suffixes, mesoclisis operates inversely, attaching clitic suffixes to an infinitival stem *before* all ordinary suffixes.

We shall return to this analysis in Chapters 8 and 9, and examine how the theory of Paradigm Function Morphology can accommodate the data about cluster formation and the preverbal placement of clitic clusters. However, before considering those aspects, I shall first investigate an inflectional analysis of the morphophonological effects induced by postverbal clitics.

Chapter 7 Morphophonological effects

As alluded to in Chapter 4, one of the distinctive properties of postverbal clitics in EP is the fact that they undergo shape alternations and induce allomorphic variation on the adjacent verbal stem. It will be assumed that these effects, which are often determined by specific person and number features of clitics and verbs, are best analysed as morphophonological phenomena.

The main goal of this chapter will be to account for these clitic-induced alternations within Paradigm-Function Morphology (Stump 2001) and to show that an inflectional analyses captures the morphophonological nature of the data more insightfully than previous accounts within phrasal phonology (Gerlach 2001a) and Precompiled Phonology (Vigário 1999b). This chapter is organised as follows: section 7.1 offers a summary of the allomorphic effects under discussion and section 7.2 surveys previous phonological accounts of the data. An inflectional analysis of pronominal allomorphy and stem-variation is sketched in section 7.3 based on Stump's (2001) morphophonological rules and corresponding metageneralisations.

7.1 Clitic-induced allomorphy

The ability for clitics to induce or undergo allomorphy is attested in various languages. In some cases, affixal clitics affect the phonological shape of the adjacent host, as for example with the English negation marker *n't* (Zwicky&Pullum 1983) which changes the form of the preceding auxiliary (e.g., the negative form of *will* is not **willn't* but *won't*). In other cases, shape variation is suffered by the clitic itself depending on the grammatical or phonological properties of the host. In Bulgarian, for example, the phonological form of the definite determiner depends on the gender and number properties of the preceding word: it surfaces as *-te* if the host is plural, but as *-ta* if its is

feminine singular (Halpern 1995). The fact that clitics change under highly idiosyncratic circumstances supports the claim that they constitute affixes.

7.1.1 Overview of the data

This section briefly surveys the various allomorphic changes triggered by clitics in EP (cf. also Cunha&Cintra 1987, Spencer 1992, Crysmann 1997, Luís 2003b). As discussed earlier (cf. Chapter 4), EP clitic suffixes trigger phonological variation on the preceding verb and themselves undergo shape variation. The phenomena are determined mostly by specific grammatical and lexical factors and can therefore not be derived through productive rules of phonology.

7.1.1.1 Clitic allomorphs

We shall start with 3rd person accusative pronouns. They can appear in one of the three forms given in (1), depending on the phonological/grammatical form of the preceding verb.

(1) 3rd person accusative (en)clitics

	3.SG.M.ACC	3.SG.F.ACC	3.PL.M.ACC	3.PL.F.ACC
default form	o	a	os	as
l- form	-lo	-la	-los	-las
n- form	-no	-na	-nos	-nas

The table in (2) illustrates the occurrence of the default form. It shows that if the preceding verb is vowel-final, 3rd person accusative clitics surface as vowel-initial forms (cf. Chapter 4 for examples).

(2) verb-enclitic construction with vowel-final forms of *dar* 'give'

	PRESENT	PERFECT	IMPERFECT	SUBJUNCTIVE
1sg	dou → dou-**o**	dei → dei-**a**	dava → dava-**o**	desse → desse-**as**
2sg	----------	deste → deste-**o**	----------	----------

3sg	dá → dá-**os**	deu → deu-**as**	dava → dava-**o**	desse → desse-**a**

If the preceding verb ends in one of the following consonants: -r, -s or –z, 3rd accusative clitics surface as *-lo*, *-la*, *-los*, *-las*. An example of consonant-final verbs is given in (3-5).

(3) verb-enclitic construction with s-final forms of *dar* 'give'

	PRESENT	PERFECT	IMPERFECT	SUBJUNCTIVE
2sg	dás → dá-**las**	----------	davas → dava-**la**	desses → desse-**lo**
1pl	damos → damo-**lo**	demos → demo-**lo**	davamos → davamo-**lo**	dessemos → dessemo-**los**

(4) verb-enclitic construction with z-final forms of *fazer* 'do'

	PRESENT	PERFECT
1sg	----------	fiz → fi-**lo**
3sg	faz → fá-**lo**	fez → fê-**la**

It is worth noting that the same type of allomorphy is triggered verb-internally when clitic suffixes appear after the infinitival stem which ends in –r, as illustrated in (5).

(5) verb-enclitic construction with r-final forms of *dar* 'give'

	INFL INFINITIV	INFINITIV	MESOCLITIC FUTURE
1sg	dar → dá-**lo**	dar → dá-**la**	darei → dá-**las**-ei
	----------		darás → dá-**las**-ás
3sg	dar → dá-**lo**		dará → dá-**la**-á

Finally, 3rd accusatives surface as *-no, -na, -nos, -nas* when preceded by 3rd plural verb forms:

(6) verb-enclitic construction with 3rd pl verbs (*pôr* 'put')

	PRESENT	PERFECT	IMPERFECT	INFL. INFINITIVE	IMPERATIVE

3pl	põem-o → põem-**no**	põem-o→ põem-**no**	põem-o→ põem-**na**	põem-o→ põem-**nos**	põem-o → põem-**nas**

Alternatively, the more conservative varieties of Portuguese (generally Central and Northern varieties) select the *n*-form after any nasal-final verb, regardless of agreement features. In that case, a 2nd person singular singular forms such as *põe* 's/he puts' or *tem* 's/he has' would also trigger the *n*-allomorph (cf. Chapter 4 for phonetic transcriptions).

Recapitulating, the contexts affecting the phonological shape of 3rd person accusative enclitics are as follows,

- the *n*-form is triggered either solely by 3rd person plural verb forms or by any nasal final verb (cf.6);
- the *l*-form of the 3rd person accusative clitic is triggered by consonant final verbs ending in -s,-z,-r (cf. 3-5);
- the vowel initial allomorph is selected elsewhere (cf 2).

7.1.1.2 Clitic-induced stem allomorphy

Having addressed the phonological changes suffered by postverbal clitic, I will now summarise the variations suffered by the verbal host.

a) *l*-induced stem allomorphy

There are two types of clitic-induced stem-allomorphs: one is triggered by 3rd person accusative *l*-forms; the other is triggered by 1st and 2nd person plural enclitics. The first case is exemplified in (7-9), where the *-lo*,*-la*,*-los*,*-las* forms induce word-final consonant deletion on verbs ending in -s, -z or –r, both verb-finally and verb-internally (cf. 9). (Consonant deletion is signalled with '∅').

(7) deletion of verb-final -s on verb forms of *dar* 'give'

	PRESENT	PERFECT	IMPERFECT	SUBJUNCTIVE
2sg	dá**s** → dá∅-las	----------	dava**s** → dava∅-la	desse**s** → desse∅-lo

1pl	damo**s** → damo∅-lo	demo**s** → demo∅-lo	davamo**s** → davamo∅-lo	dessemo**s** → dessemo∅-los

(8) deletion of verb-final -z on verb forms of *fazer* 'do'

	PRESENT	PERFECT
1sg	----------	fi**z** → fi∅-lo
3sg	fa**z** → fá∅-lo	fe**z** → fê∅-la

(9) deletion of verb-final -r on verb forms of *dar* 'give'

	INFL INFINITIV	INFINITIV	FUTURE
1sg	da**r** → dá∅-lo	da**r** → dá∅-la	da**r**ei → dá∅-las-ei
2sg	----------		da**r**ás → dá∅-las-ás
3sg	da**r** → dá∅-lo		da**r**á → dá∅-la-á

An interesting aspect of this type of stem-allomorphy is that it is associated with clitic-allomorphy. In other words, the set of consonants that trigger the *l*-form of the accusative clitics in (3-5) is exactly identical to the set of contexts that undergo deletion in (7-9). This mutual or reciprocal conditioning, as shall be argued below, poses serious problems to accounts which derive the data through phrasal phonology.

b) *-nos* and *-vos* & 1st plural verb forms

The other idiosyncratic case of stem allomorphy is triggered by the 1st and 2nd person plural pronouns *-nos* and *-vos* when they appear after a 1st pl verb forms with the agreement suffix *-mos*. In this case, the final -s of the suffix is deleted:

(10) Nós vêmo∅-nos hoje. (not *vêmos-nos).
we see-1.PL.REFL today
'We see each other today'

Further examples of consonant-deletion induced by *-nos* and *-vos* are given below:

(11) deletion of verb-final -s on 1st person plural verb forms of *dar* 'give'

	PRESIND	IMPERFECT	SUBJUNCTIVE
2sg	dás →dás-nos	davas → davas-nos	desses → desses-vos
1pl	damo**s** → damo∅-vos	davamo**s**→ davamo∅-nos	dessemo**s** → dessemo∅-vos

The crucial point about this type of allomorphy is that it is grammatically conditioned by both the agreement features of the verb and the person/number features of the clitic. Again, this type of data relies heavily on specific grammatical properties which cannot be accessed outside morphology.

To sum up then, there are two types of phonological alternations on the verb:

a) verb-final consonant deletion before *l*-accusative enclitic; and
b) verb-final consonant deletion on 1st person plural verb forms before 1st and 2nd person plural enclitics (i.e., *nos* and *vos*).

c) clitic clusters

Finally, let us also consider cluster-internal allomorphy. There are two types: a) between *–nos* and *-vos* followed by 3rd person accusative clitics (12) and b) between *me*, *te*, *lhe* and *lhes* followed by 3rd person accusative clitics (13). Both types then occur within the dative-accusative cluster (cf. Chapter 4)

(12) a. *nos-os → no-los
1pl.dat-3pl.masc.acc

b. *vos-o → vo-lo
2pl.dat-3sg.masc.acc

(13) a. me-o → mo
1sg.dat-3sg.masc.acc

b. lhe-os → lhos
3sg.dat-3pl.masc.acc

In (12), we observe a case of mutual or reciprocal allomorphy whereby the consonant-final dative clitic selects the *l*-form of the accusative clitic and the l-form triggers consonant-deletion on the preceding dative clitic. The factors determining allomorphy are exactly identical to the ones identified in (2-5) and (7-9). In (13), we have a case of vowel deletion of the dative clitic and a 'fusion' of two monosyllabic affixes.

7.1.2 Affixal evidence

One of the crucial aspects about the data shown above is that phonological alternations are determined by categorial and grammatical factors, taking place between enclitics and verbs under specific feature combinations. The properties presented in the above section clearly suggest that an inflectional analysis of EP pronominals should be preferred (Spencer 1992, Crysmann 1997, Luís to appear).

None of the variations then can be found outside the various contexts described above. For example, even though definite articles in Portuguese share with the default form of 3rd accusative clitics the same phonological form, they never undergo allomorphy. So, none of the Portuguese definite determiners *o*, *a*, *os*, *as* ever change into *lo*, *la*, *los*, *las* when preceded by consonant-final words (14a), nor into *no*, *na*, *nos*, *nas* when preceded by a nasal-final word (14b) (cf. ch.4 for further exemplification):

(14) a. todos os carros → todos *los carros
b. o João os viu → o João *nos viu

Likewise, consonant-deletion takes place in a very restricted context: it is triggered by *l*-forms of 3rd person accusatives and by 1st/2nd pl forms -*vos*/-*nos* in post-verbal position. Thus, if *nos* or *vos*

appear before the verb and after any other word form, deletion will be blocked (e.g., todos nos vêem → *todo∅ nos vêem).

In fact, *-nos* and *-vos* can only delete a verb-final -s if it is part of the 1st pl verb form. The 2nd sg agreement marker -s, for example, fails to undergo deletion when followed by a 1st or 2nd pl enclitic:

(15) Tu recebes-nos. → *tu recebe∅-nos
you.SG receive-2.SG.ACC
'You receive us'

The conclusion we can draw from the data is that the above phenomena do not constitute general phonological processes which extend to any word form of Portuguese. It is far from clear how these allomorphic effects can be accounted for by phonological rules because any account will have to refer to both the category as well as the grammatical features of the intervening units (i.e., verbs and enclitics). Particularly difficult to motivate under a phrasal phonology analysis would be the case of 'reciprocal' allomorphy where both the enclitic and the verb affect each other's shape. We shall address further problems raised by non-inflectional analysis in the following section.

7.2 Earlier accounts

Two recent approaches to Portuguese cliticisation have been presented by Vigário (1999b) and Gerlach (2001a). As alluded to before (cf. Chapter 3), these studies argue against the affixal status of clitics and treat clitics as special function words. While Vigário (1999b) treats shape variation as phrasal allomorphy, and Gerlach (2001a) derives these effects through clitic-specific phonological constraints. I will address each study in turn and argue that clitic-induced variation in EP fall neither within the scope of phrasal phonology nor of precompiled phonology. We start with the work of Vigário (1999b).

7.2.1 Phrasal allomorphy

Despite going to great lengths to challenge the morphological attachment of enclitics (cf. Chapters 4 and 5 for discussion), Vigário (1999b) supports the view that the phonological changes suffered by enclitics cannot be derived through standard phrasal phonology. To accommodate the claim that clitics exhibit non-productive shape variation, it is suggested that clitic pronouns have allomorphic variants in the lexicon. Adopting insights from Precompiled Phonology (Hayes 1990), this means that both the *l*-forms and *n*-forms of accusative pronouns are derived as phrasal (i.e., word-level) allomorphs and inserted in the syntax postlexically.

In what follows, I will briefly survey the data originally analysed by Hayes to motivate the need for 'precompiled' level of phonology (7.2.1.1). This is followed by a detailed discussion of the problems raised by Vigário's analysis of EP. It is shown that the EP data is substantially different from the data analysed by Hayes. Crucially, because for EP it is not enough to determine the context within which a given allomorph is inserted; it is necessary to account for the fact that clitic allomorphs also affect the verbal host. To capture the fact that phonological alternations affect both clitics and the preceding verb, Vigário is forced to introduce significant extensions to the framework developed by Hayes (1990), seriously challenging the spirit of the theory of Precompiled Phonology.

7.2.1.1 Preliminaries

To provide a background against which to compare Vigário's proposal for EP, I will first provide a brief summary of the phenomena that motivated Hayes' (1990) theory of Precompiled Phonology. The data used by Hayes shows that some phonological alternations are dependent on purely syntactic information. Two phonological phenomena which seem to have direct access to syntax are a) the alternation of the feminine article *la* in Spanish and b) the rule of vowel shortening in Hausa.

In the first case, *la* becomes *el* before nouns beginning in stressed /á/; in the second case, verb-final long vowels are shortened when they are preceded by a full NP direct object.

Assuming that phonology cannot have direct access to the syntax, Hayes (1990) develops a framework in which word classes may have allomorphic variants. Crucial for this proposal is the fact that each one of the phrasal allomorphs is accompanied by a frame indicating where it may be inserted in the syntax[49]. In Spanish, for example, it is assumed that the lexical entry for the feminine article contains two allomorphs, i.e. *la* and *el*, and two environments for syntactic insertion, as in (16). Under the Elsewhere Condition, the insertion context of *el* is more specific than that of *la* and therefore takes precedence over the more general.

(16) /el/ / __ [$_N$ á
/la/ / (elsewhere)

For Hausa, Hayes derives phrasal allomorphy by lexical rule. The choice of a rule rather than of lexically listed forms is based on the grounds that the phonological change of vowel shortening affects whole classes of words. The lexical phonological rule in (17a) derives the verb forms and the phonological instantiation frame in (17b) inserts these forms into their relevant syntactic context:

(17) a. V: → V / [... __] $_{[Frame1]}$
b. Frame 1: / [$_{VP}$ __ NP ...], NP non-pronominal

What is striking about the data analysed by Hayes (1990) is the fact that is significantly different from the morphophonological effects displayed by EP:

[49] It is important to point out that precompiled phonology assumes a late insertion model of syntax. Words are represented through abstract markers and phonological instantiation takes place post-syntactically (Hayes 1990).

a) in EP, phonological changes are reciprocal given that they affect both the selected pronominal allomorph *and* its context of insertion (e.g., 3rd accusative clitics surface as *l*-allomorphs and trigger consonant deletion on the verb);

b) the context within which allomorphy takes place must make reference to the specific inflectional features of the elements involved, (e.g., only 3rd plural verbs trigger *n*-form on 3rd accusative clitics, and *-nos* and *-vos* only affect 1st plural verb forms)

c) EP pronouns lack the syntactic freedom of the Spanish and Hausa word categories.

7.2.2.2 Precompiled allomorphs

This section presents the precompiled account by Vigário (1999b) and discusses the problems that emerge from a postlexical view of the verb-enclitic combination.

As mentioned before, Vigário adopts a prosodic (postlexical) view of cliticisation based on the assumption that clitic pronouns are function words which combine with the verb in the syntax (cf. Peperkamp 1997, for Spanish and Italian). Within the theory of Precompiled Phonology (Hayes 1990), she argues that clitic variants are derived as phrasal allomorphs (i.e. syntactically inserted word-level allomorphs). The leading idea of her account is that words can have allomorphic variants in the lexicon.

Let us now look at the lexical rules and insertion contexts set up by Vigário. To account for pronominal allomorphy, Vigário produces a lexical entry for 3rd accusatives where each allomorph is individually specified for its context of insertion (18).

(18) a. no / [...] Vb $_{[3pl]}$ __ (e.g davam 'they gave')
b. lo / [...] Vb$_{[+cons]}$ __ (e.g davas 'you.sg gave')
c. o / elsewhere (e.g dava 'I/he/she gave')

The first rules inserts *n*-allomorphs in the phrasal context of a 3rd person plural verb form; the second rule inserts *l*- allomorphs after a consonant-final verb; the thrird rule inserts the default form.

So far we have seen how each accusative allomorph is derived. Let us now examine how stem allomorphy is derived. In (19), consonant deletion on verbs is captured through a lexical rule which is triggered by the contexts specified in (20).

(19) C $\rightarrow \varnothing$ / [... _ ...] [Frame 1] [Frame 2]

(20) a. Frame 1: ... [..._] Vb [...] cl [acc]

b. Frame 2: ... [..._] Vb [1pl] [...] CL (nos/vos)

In (20a), frame 1 deletes the consonant before an *l*-accusative, and in (20b) frame 2 deletes the verb-final consonant of a 1st plural verb form followed by 1st or 2nd plural enclitics.

a) grammatical features

An important detail about Hayes' analysis is that allomorphs are strictly dependent on syntactic and phonological information. For EP, that is clearly not the case.

As formulated in (18a), all that appears to be necessary to select an *n*-form is the right context of insertion. The problem is that the so-called context is not just based on phonological or syntactic properties, but crucially on inflectional features of the verb form: the *-no*, *-na*, *-nos*, *-nas* allomorphs must follow verb forms with 3pl features, as Vigário points out.

Similarly, the context of insertion in (20b) is defined on the basis of very specific inflectional properties. Although the rules in (20) seem to draw an analogy between the placement of Hausa verbs and EP verbs, the context of insertion for Hausa verbs is purely syntactic and phonological, for EP it must be defined in terms of a restricted set of pronominal person and number features.

A similar problem is raised by the rule deriving consonant-less verb forms in (19). This rule is formulated in analogy with the rule of Vowel Shortening for Hausa verbs. However, the Hausa rule applies to a whole class of words, namely verbs with long vowels (17a), while in EP the effect of the

rule deriving consonant-deletion is dependent in some cases on the grammatical features of the verb, which must be a 1st person plural form.

The fact that shape variation cannot be derived by merely addressing purely phonological or syntactic contexts suggests that they do not constitute the set of phenomena precompilation was designed to account for. The idea that phonological effects motivated by morphosyntactic features should be regarded as evidence for morphological status is also pointed out in Hayes (1990), in his discussion about the distinction between true morphological alternations and phrasal allomorphs. He states that 'rules of [inflectional] allomorphy have diverse structural conditions: they may refer to phonological environment, *to inflectional features*, and to the identity of individual morphemes (...)' (p. 90, my emphasis). What we have in EP is precisely a set of clear 'inflectional features'.

b) 'reciprocal' allomorphy

One further set of problems is raised by the verb-enclitic interaction. Under Vigário's analysis, *l*-allomorphs have the ability to affect the phonological shape of their context of insertion given that they trigger deletion on consonant-final verbs. The first observation that seems appropriate is that the bi-directionality of this allomorphy is quite distinct form Hayes' Spanish and Hausa examples. In these languages, phrasal allomorphs are either listed or derived through a lexical rule but the 'frames' within which the Spanish article or the Hausa verbs are inserted are left unaffected.

Unfortunately not much is said about how Precompiled Phonology would derive the mutual allomorphy. In particular, it is not clear how the lexical entries in (18) and the rules in (19-20) would interact in order to produced the correct results. If the clitic allomorph is selected because there is a consonant-final verb in the syntax (note that insertion frames are syntactic), then it is difficult to see how the deletion rule in (19) would apply in the lexicon prior to the insertion of the verb.

Even if an analysis could be provided, substantial extensions would have to be introduced to the theory. For example, regardless of how consonant-deletion on the verb is formalised, it would be

necessary to allow that rule to have a 'destroying' effect on the syntactic context within which the *l*-allomorph has been inserted. In other words, the interaction between (18b) and (19)-(20a) entails that the phonological form of the verb can be altered after the insertion of the clitic in the syntax. It seems that deriving these reciprocal effects in phrasal phonology results in an unnatural analysis.

Among other extensions, it would also be necessary to stipulate that the rules provided by Vigário would apply in a predetermined order so as to ensure that selection of the *l*-allomorph takes precedence over verb-final consonant deletion. Of course, under an inflectional analysis, this order falls out naturally given that deletion of the consonant is part of the evaluation of the rule deriving the clitic allomorph (cf. 7.3).

c) marked status of enclitics

Other problems with a 'precompiled' analysis include the fact that enclitics are systematically treated as exceptional units, in contrast with preverbal clitics (which, for Vigário, constitute well-behaved function words). In other words, by deriving enclisis through special mechanisms such as phrasal allomorphy, it appears that enclitics are grammatically marked. Studies on acquisition of clitic pronouns however have revealed that enclitics constitute the unmarked form, found very early in child language. In adult language it is also recurrently found in contexts where proclisis would be expected (cf. section ch 3, Duarte et al.).

Related to the highly marked status of accusative allomorphs is also the fact that the contexts proposed by Vigário are not relevant for any other phenomenon. On the contrary, Hayes' theory predicts that the syntactic contexts proposed for allomorph insertion also play a role in accounting for other phenomena (e.g. Hausa). However such evidence cannot be provided for EP.

d) opaque clusters

Finally, Vigário's precompiled analysis also assumes that clitic clusters correspond to lexically listed units rather than to clitic sequences. There are two serious problems with this assumption. First, there is the general insight that viewing clusters simply as opaque units inevitably overlooks the transparent nature of most clitic sequences and cluster-internal regularities. From a prosodic/syntactic point of view, deriving clusters as combinations of individual clitics poses serious difficulties given that it is quite difficult to capture the various idiosyncrasies which characterise clitic combinations such as co-occurrence restrictions and clitic order (Anderson 1995); and therefore an 'opaque' view of clusters may turn out to simplify the analysis. However it evidently leaves an important aspect of cliticisation unaccounted for. Failure to take into account the internal composition of clusters seriously weakens any analysis. Second, by treating clusters as opaque units Vigário's analysis fails to capture the fact that consonant-final deletion triggered by *l*-forms of 3rd accusative clitics also takes place inside clitic clusters, as in *no-lo* and *vo-lo*, for example (cf. 12). As I shall illustrate in section 7.3, any overall analysis of the data should be able to naturally capture the regularity underlying this clitic-induced deletion.

To conclude, there are various conceptual and empirical problems which indicate that the EP data should not be analysed within Precompiled Phonology. The next section briefly addresses Gerlach's main claims about the phonological alternations.

7.2.2 Clitic-specific phonology

Gerlach (2001a, 2001b) draws on clitic pronouns in various Romance languages, including EP. In analogy with Vigário (1999b), it is argued that Romance clitics do not constitute affixal units. But while Vigário tries to assimilate clitics to the category of function words, Gerlach assumes that they form a theoretical category of their own (cf. Chapter 3). By assigning theoretical status to clitics, Gerlach introduces clitic-specific phonological constraints for the derivation of clitic-induced shape

alternations. More precisely, it is claimed that the underlying force driving the allomorphy is essentially phonotactic, being determined by syllable structure and of prosodic constraints of the type NO HIATUS, NO CODA and ONSET. I will show next that these constraints only account partially for the data.

a) ONSET

The tendency for vowel-initial enclitics to trigger consonant-deletion on the verbal host is accounted for by Gerlach (2001a, 2001b) by assuming that EP gives primacy to onsets over codas. So, for example, the *l*-consonant of the 3rd accusative allomorph in enclitics appearing in the onset provides the enclitic with an onset position at the expense of the coda of the preceding syllable (syllables are signalled with '.'):

(21) a. *ven.des.-o → ven.de.lo
'(you.sg) sell it'
b. *com.prar.-o → com.pra.lo
' (to) buy it'

This generalisation applies, in effect, nicely to (21). But it would fail to explain the onset-less clitic in (22a). As mentioned earlier, some varieties of EP place the *n*-allomorph only when the nasal diphthong corresponds to the 3rd pl marker as in (22b) (Vigário 1999b). If it is indeed the case that the language prefers CV syllables, then the vowel-initial clitic in (22a) is left unexplained:

(22) a. põe-as
'(s/he) puts them'
b. põem-no
'(they) put it'

b) NO HIATUS

Adjacent vowels in EP can be avoided in various ways. One of them is illustrated in (23) where the the marker for conjugation class, tense and agreement, namely –e, is semivocalised before a 3rd person accusative enclitic (cf. 5.2.2.4)

(23) a. me-te-a → me.t[ju]
's/he inserts it'
b. be-be-o → be.b[ju]
's/he drinks it'

Note however that the cliticised verb forms in (23') also contain two adjacent vowels but none of them is semivocalised. The hiatus in verb forms with stressed final vowels is simply not broken up, neither through glide insertion nor through assimilation of the vowels.

(23') a. vê-as
's/he sees them'
b. dá-as
's/he gives them'
c. lê-as [
's/he reads them'

c) NO CODA

One further case of shape variation which fails to support Gerlach's claim about syllable structure is the occurrence of consonant deletion before 1st and 2nd person plural enclitics. Even though it might at first seem that deletion is triggered by a constraint which penalises syllables with codas, this view does not explain why only 1st person plural verbs undergo deletion. In (24), the last syllable before the enclitic is phonotactically identical, but only (24a) undergoes deletion.

(24) a. es.cre.ve.mo-vos → *es.cre.ve.mos-vos
'(we) write to you'

b. es.cre.ves-nos → *es.cre.ve-nos
'(you.sg) write to us'

To sum up, the idea that clitic-induced shape variation is determined by phonotactic constraints provides interesting generalisations about the forces driving some of alternations, but they do not account for the fact that most of these changes have grammaticalised and are best accounted for by making reference to grammatical features. I therefore conclude that it is not only possible but also desirable to to derive the shape alternations as morphophonological effects within inflectional morphology.

7.2.3 Summary

Previous accounts of clitic-allomorphy in EP have derived the phenomena through 'special' phrasal phonology (Vigário 1999b) or clitic-specific phonotactic constraints (Gerlach 2001a). Both studies share the idea that clitics do not constitute affixes and that the observed shape variation can be accounted for through fundamentally phonological generalisations. The problem with these analyses, as I tried to show, is that clitic-allomorphy in EP is induced by specific word-categories and also by specific grammatical features; both these factors are typical of grammatically conditioned morphophonology (Matthews 1991, Stump 2001, Anderson 1992, Spencer 1992, Aronoff 1995).

7.3 Morphophonological rules

This section sketches a purely inflectional analysis of clitic induced allomorphy in EP, adopting Stump (2001)'s treatment of inflectional allomorphy. Section 7.3.1 introduces the PF-formalism for the treatment of allomorphy and section 7.3.2 illustrates how such an analysis might look.

7.3.1 The formalism

Certain allomorphic effects are triggered by a specific group of affixes. Trisyllabic laxing in English, for example, is induced by the suffixation of *-ity* (as in div[aj]ne vs div[I]nity) but not with the suffixation of *-able*. To express this association, it is widely assumed that inflectional rules should be associated with a particular class of morphophonological rules (Kiparsku 1982, Siegel 1979, Anderson 1992). In Paradigm-Function Morphology, in particular, associations between inflectional rules and morphophonological regularities are captured through morphophonological rules that are part of the realisation rules themselves (Zwicky 1992, Bochner 1993, Orgun&Inkelas 1998).

To illustrate how RRs interact with morphophonological (m-p) rules, let us consider an example from Bulgarian verbal morphology (Stump 2001, Chapter 2). One of the realisation rules suggested by Stump for the derivation of verbal agreement in Bulgarian is the rule in (25) realising the properties 'third person plural present tense'.

(25) Realisation rule for Bulgarian 3pl agreement marker

$RR_{E8,\ \{Tns:\ pres,\ Per:3,\ Num:\ Pl\},\ V}$ (<X, σ>) = $_{def}$ <Xɵt', σ> (Stump 2001:45)

What is interesting about this rule is that it it causes a stem-final vowel to delete (e.g., *igráje* 'play' is derived as *igrájɵt*), and it also causes the suffix to be stressed if the elided stem-vowel is stressed

(e.g., *kradé* 'steal' is derived as *kradə́t*). The morphophonological rule capturing this regularity is formulated as in (26):

(26) Morphophonological rule for fragment of Bulgarian RRs (Stump 2001:48)

"If X = W [vowel$_1$] and Y = X [vowel$_2$]Z, then the indicated [vowel$_1$] is absent from Y' and the indicated [vowel$_2$] is stressed in Y' iff [vowel$_1$] is stressed in Y."

To ensure that the rule in (26) is associated with the correct realization rules, Stump provides morphophonological metageneralisations which aim at constraining the application of realization rules. In the case of Bulgarian, it is assumed that the rule in (26) only applies to a specific set of inflectional rules (cf. Stump 2001, (18) on p. 49-50). If two or more rules exhibit the same morphophonological regularity, then morphophonological metageneralisations simply express that a given class of realization rules behave in the same way with respect to some morphophonological rule (Stump 2001, p.47).

Summarising then, Stump's treatment of allomorphy entails that morphophonological rules and realisation rules are hierarchically organised in the sense that the former help define the latter. Let us now examine how this proposal might be adopted for EP.

7.3.2 Proposal

Recapitulating the EP facts, (27) summarises the morphophonological regularities we need to derive.

(27) Summary of morphophonological effects triggered by clitic suffixes
  a. At the verb-enclitic boundary, 3rd person accusative clitics (regardless of number or gender) can exhibit the following allomorphs:
    i. *n*-initial form after 3rd person plural verb forms (cf. 6);

ii. *l*-initial form after verbs ending in –s, -z, -r (cf. 3-5); in this context, the accusative clitic triggers stem-final consonant deletion (cf.7-9).

b. 1st and 2nd person plural *-nos* and *-vos* trigger stem deletion on 1pl verb forms (cf. 11).

c. In the dative-accusative cluster,

i. 3rd accusative clitics trigger vowel deletion on 1st sg, 2nd sg and 3rd sg dative clitics (cf.13a,b)

ii. 3rd accusative clitics trigger vowel and consonant deletion on 3rd pl dative clitic (cf. 13c).

iii. *nos* and *vos* trigger *l*-initial accusative allomorph, followed by consonant deletion of final –s in *nos*/*vos* (cf.12).

Of the alternations summarised above, (27a) and (27b) describe effects which take place at the boundary between the verb and the clitic suffix, while the effects in (27c) refer to cluster-internal phenomena. Even though the effects should be partitioned according to these two distinct context, there is one type of variation which occurs both at the verb-clitic boundary and cluster-internally, namely, the effect summarised in (27aii) and (27ciii), in which the *l*-initial allomorph is triggered by an adjacent consonant and induces its deletion.

To derive these regularities, I propose the morphophonological rules in (28) and the set of metageneralizations in (31). Each set of rules will be considered in turn.

7.3.2.1 Morphophonological rules (m-p rules)

The clitic-induced allomorphy shall be captured by the morphophonological rules given below:

(28) Morphophonological (m-p) rules for clitic-induced allomorphy in EP

Where $RR_{n, \tau, C}(X, \sigma) =_{def}$ <Y', σ>:

a. If X = Wv[nasal] and Y = XLNZ, then Y' = XnZ

b. If X = W[consonant] and Y = XLNZ, then Y' = WlZ

c. If X = WE(s) and Y = XLNZ, then Y' = WZ

d. If X = Ws and Y = XZ, then Y' = WZ

Underlying (28) is the assumption that each m-p rule constrains the evaluation of one or more realisation rules. To understand how this works it is important to recall the general format for RRs (cf. Chapters 2 and 6). What is important is the prime notation on Y' which indicates that Y' constitutes an allomorph of Y, where Y represents the default phonological form derived by a RR and where Y' is derived from the association of a given RR with a morphophonological rule. If Y is not subject to any MPR, then by default Y´= Y. Morphophonological rules therefore are not stated over forms, but determine how a given RR is evaluated.

It is also important to note that the m-p rules make reference to specific morphophonological representations of the roots and affixes whose form they affect. This idea follows Stump (2001:48)'s treatment of Bulgarian morphology in which an abstract 'morphophoneme' is postulated, namely the 'A' vowel, to capture the fact that it can surface as /e/ in certain defined contexts and as /a/ by default. For EP, one might also make use of an abstract morphophoneme for the *o*(*s*)/*a*(*s*) clitics and the *me*, *te* and *lhe*(*s*) clitics. As alluded to before, 3rd person accusative clitics have three allomorphs: l-type, n-type and the default. Let us therefore assume that these clitics begin with an abstract consonantal morphophoneme, call it 'LN', whose default realization is zero. Similarly, to account for the fact that /e/ can be realized or not, we can assume a special morphophoneme E, whose default value is /e/. Given this assumption, the clitic block given in Chapter 6 will need to be slightly rewritten. The change is illustrated below for rule block C and some forms of Block B:

(29) RRs for 3rd person accusative clitic suffixes (Block C)

a. $RR_{C1,\ \{Case:\ Acc,\ Person:3,\ Number:\ Sg,\ Gender:\ Masc\},\ V}$ (<X,σ>) $=_{def}$ <X-LN*o*, σ>

b. $RR_{C2,\ \{Case:\ Acc,\ Person:3,\ Number:\ Sg,\ Gender:\ Fem\},\ V}$ (<X,σ>) $=_{def}$ <X-LN*a*, σ>

c. $RR_{C3,\ \{Case:\ Acc,\ Person:3,\ Number:\ Pl,\ Gender:\ Masc\},\ V}$ (<X,σ>) $=_{def}$ <X-LN*os*, σ>

d. $RR_{C4,\ \{Case:\ Acc,\ Person:3,\ Number:\ Pl,\ Gender:\ Fem\},\ V}$ (<X,σ>) $=_{def}$ <X-LN*as*, σ>

(30) RRs for 1st and 2nd person and 3rd dative clitic suffixes (Block B)

a. $RR_{B1,\ \{Person:1,\ Number:\ Sg\},\ V}$ (<X,σ>) $=_{def}$ <X-*mE*, σ>

b. $RR_{B2,\ \{Person:2,\ Number:\ Sg\},\ V}$ (<X,σ>) $=_{def}$ <X-*tE*, σ>

c. $RR_{B3,\ \{Person:1,\ Number:\ Pl\},\ V}$ (<X,σ>) $=_{def}$ <X-*nos*, σ>

d. $RR_{B4,\ \{Person:2,\ Number:\ Pl\},\ V}$ (<X,σ>) $=_{def}$ <X-*vos*, σ>

e. $RR_{B5,\ \{Case:\ Dat,\ Person:3,\ Number:\ Sg\},\ V}$ (<X,σ>) $=_{def}$ <X-*lhE*, σ>

f. $RR_{B6,\ \{Case:\ Dat,\ Person:3,\ Number:\ Pl\},\ V}$ (<X,σ>) $=_{def}$ <X-*lhEs*, σ>

Let us now consider the m-p rules in (28) in more detail: X represents the base which serves as input to a given RR; Y constitutes the output of applying RR to a base prior to the application of any m-p rule, and Y' gives us the appropriate allomorph of Y. Given this brief notational clarification, let us consider the generalisation encapsulated by each m-p rule.

The rule in (28a) says that if X is a verb form with a final nasal diphthong and if this verb form is followed by an accusative clitic, then the output of combining the verb with the clitic will trigger one of the following *n*-initial allomorphs: *-no(s)*, *-na(s)*. Illustrating, if the verb form is *levam-* and the clitic is 3rd person singular masculine, then the output will be *levam-no* (cf. 27ai).

According to m-p rule (28b) if a given base is consonant-final and if it is followed by one of the 3rd person accusative clitics, the *l*-form of the accusative clitics must surface (cf. 27aii and 27ciii). It also says that the selection of the accusative allomorph affects the phonological shape of the preceding base by triggering deletion of the verb-final consonant. This m-p rule then expresses both the derivation of the *l*-allomorph and the consonant deletion on the preceding base. Note that this rule applies both at the verb-enclitic boundary, as in *dava-lo*, and at the clitic-clitic boundary, as in *no-lo* or *vo-lo*. So, for example, if the base to which the clitic *-a* attaches is *lavas-* then the output will be *lava-la*; if the base is *levas-nos-* then the output Y' will be *levas-no-la*.

Rule (28c) derives the 'fused' dative-accusative cluster by deleting the ending on the dative clitic (cf. 27ci-cii). So, if for example we want to combine the base *levam-me-* with the clitic -o, the output Y' will be *levam-mo*.

Rule (28d) derives the stem-allomorph that is triggered by 1st and 2nd person clitics on 1st plural verbs (cf. 27b). The rules causes a consonant final stem bearing the features {Person:1, Number: Pl} to be replaced by a consonant-less allomorph. So, if for example X is *levamos-* the Y' will be *levamo-nos*.

7.3.2.2 Morphophonological metageneralisations

Given that m-p rules cannot be associated to all realisation rules of EP, we need metageneralisations to capture the valid associations between m-p rules and realisation rules.

(31) Metageneralisations associated with the m-p rules in (28)

a. m-p rule (28a) is associated with realisation Clitic-Block C, if σ_1 contains {Agr:3pl};
b. m-p rule (28b) and (28c) are associated with Clitic-Block C, everywhere.
c. m-p rule (28d) is associated with the RR_{B3} and RR_{B4} from Clitic-Block B, if σ_1 contains {Agr:1pl}

The metageneralisation in (31a) says is that the m-p rule deriving *n*-allomorphs applies only to the realisation rules deriving 3rd accusative clitics (cf. 29). In addition, the m-p rule only applies within a specific grammatical context: if the clitic is preceded by 3rd person plural verb forms. For the more conservative dialects of EP in which any verb ending in nasal diphthong can select an *n*-allomorph (cf. discussion in Chapter 4 and above in 7.1), we simply need to relax the featural constraint.

(31b) associates the m-p rules (28b) and (28c) with 3rd person accusative clitics. For these m-p rules no further stipulation is necessary given that the m-p rule itself provides the necessary information: as alluded to before, (28b) is phonologically conditioned by a consonant-final base which can be either a verb or a clitic; the m-p rule (28c) applies if 3rd accusatives occur after the clitic forms mE, tE and lhE(s).

Finally, (31c) associates the m-p rule (28d) with two specific realisation rules from Block B, namely RR_{B3} and RR_{B4} deriving the clitics *–nos* and *–vos*, as shown in (30). This m-p rule however is also grammatically conditioned by the person and number features of the verb preceding the clitic. We therefore need to restrict the application of the m-p rule to a very specific grammatical context, namely to 1st plural verb forms. From the metageneralisations in (31) it is clear that not all realisation rules deriving clitics are associated to m-p rules.

What is interesting about this proposal is that it captures shape variation by enriching the inflectional rules with information about the grammatical and phonological contexts that influence the realisation of their affixes. In addition, the account is also parsimonious given that the rules apply to several RRs (in this case, (28a,b) condition the evaluation of the RRs in (29), instead of listing all 3rd person accusative allomorphs (cf, 7.2) which misses the generalisation that they all undergo the same type of shape variation.

7.4 Summary

The aim of this section has been to illustrate how clitic-related shape variations in EP can be derived as part of inflectional allomorphy. This account is superior to Gerlach (2001a) in that it refers not only to phonological properties of the data but also to the lexical and grammatical factors determining variation. In addition, the use of morphology-internal mechanisms captures quite naturally the role played by specific featural combinations.

With this chapter, then, we have come to the end of the second part of this thesis. We have so far given emphasis to suffixing clitics (i.e. enclitics and mesoclitics) in EP and argued that they are in morphological construction with the verb. Recapitulating somewhat, Chapters 4 and 5 motivated the inflectional analysis by providing morphological and phonological evidence. Chapter 6 then showed that inflectional rules can successfully capture the affixal nature of clitic suffixes by deriving them

as inflectional exponents. This chapter, on the other hand, addressed some of the allomorphic alternations found at the boundary between verbs and suffixing clitics and provided a morphophonological derivation of the observed shape variations. In the third part of the thesis, the discussion shall be extended to preverbal clitics (Chapter 8) and the morphological derivation of clitic sequences (Chapter 9).

Part III – Realisation and placement of clitic clusters

Part III examines the asymmetry between preverbal and postverbal placement and investigates why the order inside the cluster remains invariant regardless of whether clusters are placed as enclitic or as proclitic sequences.

We start in CHAPTER 8 by focusing on proclitics, showing that they are formally exactly identical to enclitic clitics, although they exhibit distinct distributional and scopal behaviour. Based on this evidence, it is argued that proclitics are best regarded as *phrasal* affixes.

CHAPTER 9 aims at providing an inflectional account of the asymmetric placement of clitics and clitic clusters in EP. Revising the analysis of enclisis provided in Chapter 6, it is argued that an accurate analysis of cluster realisation and placement presupposes that clitic clusters be derived prior to their placement. It is further assumed that placement rules can determine whether the cluster combines with a verbal stem in the morphology (for enclitics) or with a phrasal node in the syntax (for proclitics).

In CHAPTER 10, I try to articulate the inflectional PF-analysis with some of the more syntactic properties of EP clitics, such as the syntactic conditioning of proclisis. It is argued that the syntactic nature of the factors determining clitic placement in EP do not invalidate the inflectional status of the clitic system. This chapter adopts a lexicalist model of syntax, within the theory of Lexical-Functional Grammar (Bresnan&Kaplan 1982, Bresnan 2001).

CHAPTER 11 summarises the main findings.

Chapter 8 Proclisis in European Portuguese

This chapter examines the behaviour of pronominal clitics in preverbal position and investigates whether proclitics are categorially different from enclitics or whether proclitics and enclitics constitute the same affixal unit.

Section 8.1 considers scopal and distributional effects displayed by proclitics which indicate that proclitics, unlike enclitics, do not form a morphologically cohering unit with the verb (Martins 1994, Crysmann 1997, Matos 2000, Vigário 1999b, Gerlach 2001a, Luís 2003a,b, Luís&Sadler 2003). Section 8.2 then argues that, despite the apparent syntactic behaviour of proclitics, they share with enclitics a wide range of significant properties which cannot be derived if proclitics are analysed as function words (as suggested by Vigário 1999b). Section 8.3 argues that proclitics constitute, in effect, phrasal affixes. Under this view, the asymmetry between enclitics and proclitics should be captured as a difference in status between the word-level and phrase-level placement of the same affixal clitic. A short discussion of the syntactic implications of the inflectional status of clitic pronouns is provided, within lexicalist assumptions about syntax.

8.1 The behaviour of proclitics

The ability of clitic pronouns in EP to appear preverbally is illustrated in (1). They occur individually in (1a) and combined into clitic sequences in (1b).

(1) a. O João não te conhece
the João not 2sg.dat know
'João doesn't know you'

b. Não se lhe deu a devida atenção.

not 3.sg.ref 3.sg.dat gave the necessary attention
'one didn't give him/her the necessary attention'

The fact that clitic sequences can also appear in preverbal position is an issue that shall concern us in the next chapter. For the moment, the goal will be to examine the distributional and scopal properties of proclitics, in particular, the fact that they can be separated from the verb by lexical items and the fact that they can have wide scope over coordinated VPs (Martins 1994, Crysmann 2002, Luís, 2002, Luís&Sadler 2003). This data, as shall be argued, challenges the wordhood status of the proclitic-verb unit.

8.1.1 Interpolation

8.1.1.1 Some data

One of the properties of proclitics is their ability to be separated from the verb by single (mostly) monosyllabic words. The most frequently found word in interpolated position is the negation marker *não* which appears both in spoken and written EP (2).

(2) a. Ela agora já me *não* quer.
She now anymore 1sg.acc not wants
'she doesn't want me anymore now'

b. Ela prometeu que lhe *não* diria nada (Barbosa 1996)
she promised that 3.sg.dat not would-say anything
'She promised that she wouldn't say anything to him/her'

Particularly relevant is the fact that interpolated elements also occur between a portmanteau cluster and the verb (3). Note that 'fused' clusters cannot be derived in the syntax, as has been widely argued in the literature (cf. section 8.2.2, for further discussion).

(3) Quando ele chegar, já lho *não* dirá.
when he arrives, anymore 3sg.dat-3sg.acc.masc not will-say
'When he arrives, s/he will not say anything about it to him/her anymore'

Other lexical items appearing in interpolated position include adverbs (4a), subject pronouns (4b) and nouns (4c).

(4) a. Eu sei que ele o *ainda* encontrará.
I know that he 3.sg.masc.acc still find
'I know that he will still find it'

b. Quantas vezes te *eu* disse para estares calado (Barbosa 1996)
how-many times 2sg.dat I told to be quiet
'How many time did I tell you to be quiet'

c. Elá só passa se a *Deus* ajudar (Barbosa 1996)
she only passes if 3sg.fem.acc God helps
'She can only pass with God's help'

As to subject pronouns and nouns, only a few can occur in that position: 1st person personal pronouns *eu* 'I' and *nós* 'we' and the noun *Deus* seems to be the most attested cases. Adverbial particles, on the contrary, are more varied: an informally collected corpus (based on adult spontaneous spoken language from Central and Northern Portugal and on informal written discourse) revealed that in addition to *ainda* (4a), it is also possible to find aspectual adverbs (5a), intensifiers (5b) and locative adverbs (5c).

(5) a. … embora eu saiba que a *já* tens em grande dose.
… although I know that 3sg.fem.acc already have in big amount
'… although I know that you already have lots of it' (e-mail message)

b. Aquilo que lhe *mais* custa a admitir é que .....
that which 3sg.dat more find-difficult to admit is that ...
'What is most difficult for him to admit is that ...' (spontaneous speech)

c. Eu nunca mais a *lá* deixava!
I never more 3sg.fem.acc there would-leave
'I would never leave it there again'(spontaneous speech)

As to the inventory of elements available for interpolation, it is not clear what exactly determines the use of some words over others. Even though interpolated particles are mostly monosyllabic, not all monosyllabic adverbials are attested. Syllable weight therefore does not seem to be the only factor determining the occurrence of these words. In effect, sometimes up to two words can co-occur, as in the frequently heard sequence *ainda não* 'not yet'. It seems instead that interpolation has evolved into a partially lexicalised phenomenon. No exact inventory of the occurring particles exists and thus further research will be necessary before exact claims are made about the circumstances under which the phenomenon takes place. Our own research revealed that emphatic structures appear to trigger the process more naturally, such as (5c). The phenomenon also appears to be subject to dialectal (Barbosa 1996) and perhaps, to some extent, idiolectal variation.

The data then shows that interpolation in contemporary Portuguese is far from being as productive as in Old Portuguese (OP) where full phrases can appear in this position (Lobo 1997, Martins 1999). As shown in (6), in OP the verb can be separated from the proclitic by a negation marker (6a), by adverbs (6b) and by the modal adverb 'assim' (6c).

(6) a. que as *nõ* quyserem (Martins 1994)
that 3sg.fem.acc not want.3spl
'they they don't want them'
b. que me *muito* pesou (Lobo 1997)
that 1sg.dat very feel-sorry
'that made me feel very sorry'
c. que lhes *assi* vemde (Martins 1994)
that 3.pl.dat then sells
'that 3sg then sells them'

It can also be separated by more than one word: the negation marker and an adverb in (7a) or the negation marker with a subject pronoun in (7b).

(7) a. se os *jaa* *nã* tiver havido em casamentos (Lobo 1997)
if 3sg.masc.acc already not had had in marriages
'if (s/he) has not already obtained them through marriages'

b. E sse o *nos* *nô* quisermos (Martins 1994)
and if 3sg.masc.acc we not want
'and if we don't want it'

In addition, we also find phrasal constituents separating the proclitic from the verb: coordinated pronominals in (8a), prepositional phrases in (8b) or a subject-object-negation sequence in (8c).

(8) a. sejã sempre como as *vós* *e* *eu* desejamos (Lobo 1997)
be always like 3pl.acc you.pl and I desire
'(that) they be always as you and I desire them'

b. quaaesquer Juramêtos que lhe *cô* *dreyto* fforem pedidos
any oaths that 3sg.dat with right could-be demanded.
'any oaths that could be demanded from him by right' (Martins 1994)

c. Se me *Deus* *enton* *a* *morte* *non* deu (Martins 1994)
if 1sg. dat God then death not gave
'if God then did not send me death'

The variety of interpolated structures seem to indicate that there is almost no restriction on the type and number of constituents that can appear in interpolated position (cf. Martins 1994).

In Modern Portuguese, on the contrary, phrasal constituents are completely ruled out, as shown in (9) where the adverbial phrases *na semana passada* 'last week' can appear before the proclitic (9a) but not between the proclitic and the verb (9b). There are therefore quite severe limits on what can be interpolated between the object cluster and verbal head (cf.2-5).

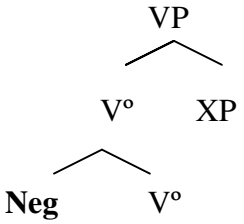
(9) a. Eu sei que a Maria *na semana passada* te telefonou.
I know that the Maria in-the week last 2sg.acc phoned
'I know that Maria phoned you last week.'

b. *Eu sei que a Maria te *na semana passada* telefonou.

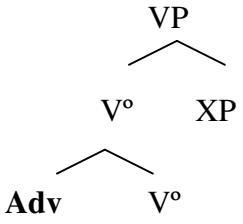
Based on the restricted separation of proclitics from the verb, I shall capture the residual nature of interpolation through the configuration in (10). In (10a-b), Neg and Adv are follows treated as non-projecting word-level units adjoined to Vº. The adjoined syntactic position is restricted to lexically marked zero-level units[50].

(10)

a)



b)



Interpolated elements can be regarded as non-projecting words that are lexically marked to appear in the interpolated position. In (10), they form with the verb a small (phrasal) construction which is dominated by the zero-level projection Vº. The 'small construction' comprises both the verb and any one of a given set of non-projecting words (e.g., the negation marker *não* and some adverbials). In addition, by treating interpolation as a lexicalised phenomenon, it is also possible to explain why the number of interpolated items can vary dialectally and idiolectally. One need only assume that the list of lexically marked items is subject to variation[51]. In this proposal, the verb appears in V, but the

[50] This proposal follows previous work on non-projecting word-level units by Sadler (1997) for Welsh and by Toivonen (2000) and Sells (2001) for Swedish, within Lexical-Functional Grammar (Kaplan& Bresnan 1982, Bresnan 2001).

[51] The phenomenon of interpolation has so far received little attention in the literature. In the present study, I have only considered non-coordinated clauses. The syntactic claims made so far are

idea could easily be recast into a IP representation of the clause, with the finite verb in I as assumed in Costa (2000), Matos (2000), and references therein.

8.1.1.2 Interpolated elements

To account for the restricted nature of interpolation, it might be argued that the elements appearing in interpolated position in modern Portuguese are morphologically attached to the verb, either through affixation or through compounding. Within this analysis, the fact that only the negation marker and a reduced number of adverbials can separate the clitic from the verb would follow from morphological restrictions. However, as the following data shows, the interpolated elements given in (2-4) behave like syntactic word units.

a) 'não'

As to the negation marker, which is by far the most productive interpolated element, it could be argued that it constitutes a verbal prefix. Several facts however indicate that it does not share any of the properties of morphologically attached elements. First, unlike affixes, it can be coordinated (11a), it can constitute an utterance on its own (11b) and it can appear in isolation (11c).

(11) a. Gostas da tua prenda? Sim e não.
like of your present? yes and no
'Do you like your present? Yes and No'

b. Sabes a resposta? Não.
know the answer? no
'Do you know the answer? No'

c. O João gostou, mas o António não.
the João liked, but the António not
'João liked it, but António didn't'

---

therefore based solely on the data presented in 9.1.1. I will leave a more detailed study of the subject for future study.

Second, there are no morphophonological effects between the verb and the negation marker. This is in clear contrast with, for example, the contracted negation marker *n't* in English which, as Zwicky&Pullum (1983) have shown, behave like an affix. The affix-like properties of English *n't* include not only the inability to appear on its own or to be coordinated, but also the fact that it triggers non-productive phonological changes on some of the auxiliary verbs it combines with (12-13). Note that the positive form of the verb in (12-13a) cannot be related to the negative form (12-13b) through regular phonological rules (Zwicky&Pullum 1983).

(12) a. *do* /du/ b. *don't* /dont/

(13) a. *will* /wil/ b.*won't* /wont/

In EP, on the contrary, the negation marker *não* displays none of these morphophonological effects: it can co-occur with all types of lexical or auxiliary verbs; it does not affect the phonological shape of the verbs; it forms with the verb a semantically transparent unit. Further evidence indicating that *não* cannot be morphologically combined with the verb is provided by the lack of arbitrary gaps. This is also in contrast with English *n't* which does not combine with all positive auxiliary forms:

(14) a. am - *amn't
b. may - *mayn´t.

So, while the combination between the auxiliary and the contracted negation marker in English cannot be derived through purely syntactic terms, the combination between *não* and verbs is completely regular. Thus, no positive evidence exists to support the affixal status of *não*. The same arguments are valid for the set of adverbial particles (such as *ainda*, *já* and *mais*) which also appear as interpolated elements but which do not show any signs of being lexically part of the verb.

b) adverbials

As to adverbials, evidence suggests that they do not form a morphological unit with the verb, unlike in Modern Greek, where compound-specific morphology clearly indicates that the adverb are part of the morphology of the verb and do not constitute independent syntactic words (Smirniotopoulos and Joseph 1998). Among the data supporting the compound-like nature of adverb-verb units are phenomena such as the union vowel *–o*. It occurs at the end of the first member of the adverb-verb complex (15b) but it fails to occur when the adverb combines freely with the verb, as in (15a):

(15)	a.	i maria	férete	*kaká*	s	tin adelfí	tis
		the Mary-nom	behaves-3sg	badly	to	the sister-acc	her
	b.	i maria		*kako*férete	s	tin adelfí	tis
		the Mary-nom		behaves-3sg badly	to	the sister-acc	her

In addition, the type of adverb-verb combinations are somewhat unpredictable (cf. 16), and the meaning of adverb-verb units can also be non-compositional (cf.17). Both these properties indicate that the relation between the verb and the adverb is lexically determined (Smirniotopoulos and Joseph 1998).

(16) a. eksijísu kalá!
Explain yourself.
b. *kaloeksijísu!

(17) a. káθome kalá
'I am behaving well'
b. kálo-káθome
'I am sitting comfortably'

On the contrary, in EP, neither the negation marker nor the adverbials display such properties. They behave like syntactic units: a) the meaning of the adverb/não-verb unit is perfectly transparent;

b) there is no compound-like allomorphic effect found at the boundary between the verb and the interpolated elements; and c) interpolation is not affected by specific classes of verbs and could - in principle – occur with any verbal lexeme. We therefore conclude that the elements appearing in interpolated position constitute lexical items.

Summing up, then, it has been shown that proclitics in EP do not have to be immediately adjacent to the verb as shown in (18), unlike enclitics in (19) which cannot be separated from the verbal host (cf. Chapters 4-7).

(18) a. O João não te conhece
the João not 2sg.dat know
'João doesn't know you'

b. que te não conheço
that 2sg.dat not know
'that (s/he) doesn't know you'

(19) a. Comprei-a *já* para ti
bought-3.sg.fem.acc already for you

b. *Comprei- *já* -a para ti
bought already 3.sg.fem.acc for you
'I bought it for you immediately'

It has been further argued that the elements which can separate the proclitic form the verb constitute word-level units, indicating that there is a syntactic position between the clitic and its host (cf. section 9.3. 2 for further discussion).

8.1.2 Coordination

One further difference between enclitics and proclitics in EP is the fact that proclitics can have wide scope over coordinated verb phrases. As pointed out by and Matos (1994, 1999) and Crysmann (1999, 2002), while narrow scope is mandatory for enclitics, it is optional for proclitics.

In a coordinated verbal structure, enclitics must appear on each verbal conjunct. As argued in Chapter 4 (cf. 4.2.2), the fact that they behave in this fashion has been used to support the idea that they constitute verbal suffixes:

(20) a. A Maria escreve-me cartas e envia-me postais.
the M. writes-1sg.dat letters and sends-1sg.dat postcards
'M. writes me letters and sends me postcards'

b. *A Maria escreve-me cartas e envia postais.
the M. writes-1sg.dat letters and sends postcards
'Maria writes me letters and sends me postcards'

In (20a), the proclitic *me* appears on each conjunct in the coordinated verb phrase *escreve cartas e envia postais*, however in (21b) it takes wide scope being shared by the conjoined VPs.

(21) a. Eu sei que a Maria me escreve cartas e me envia postais
I know that the M. 1sg.dat writes letters and 1sg.dat sends postcards
'I know that M. writes me letters and sends me postcards'

b. Eu sei que a Maria me escreve cartas e envia postais
I know that the M. 1sg.dat writes letters and sends postcards
'I know that M. writes me letters and sends me postcards'

Once again, this syntactic effect is also attested with portmanteau clusters: the clitic cluster *lho* '3.sg.dat/3.sg.acc.masc' is shared across semantically unrelated verbs in (22a) and across verb+complement phrases in (22b). We'll return to this point in section 8.2.2.

(22) a. Acho que lho comprei e ofereci.
think that 3sg.dat/3sg.masc.acc bought and gave
'I think that I bought it for her/him and gave it to her/him'

b. Nós sabemos que a Maria lho pediu de manhã
we know that the M. 3sg.dat/3sg.masc.acc requested in-the morning
e devolveu à noite.
and returned in-the evening
'We know that M. requested it (from her/him) in the morning and returned it to her in the evening.

In the other Romance languages, the ability of proclitics to be shared across coordinated structures only seems to be possible across semantically related verbs. These are generally analysed as compound-like units rather than as verb phrases (Miller 1992, Monachesi 1999):

(23) Paul les lit et relit sans cesse. (Kayne 1975)
Paul them reads and re-reads without stop
'Paul reads and rereads them incessantly'

Under the claim that *lit-et-relit* form one complex word form, the affixal behaviour of proclitics in French is not weakened (cf. Monachesi 1999, for Italian) being instead compared to that of prefixes of a compound-verb. In this respect prefixal clitics in (23) resemble the English prefixes *anti-* and *pre-* which, as shown in (24), can be shared across conjoined nouns (DiSciullo&Williams 1987).

(24) a. anti-Bush and Reagan
b. pre- and post-war

The case of EP is however more complex. Even though proclitics can also have scope over semantically related verbs, the above data has shown that they can also be shared by VPs.

8.1.3 Summary

The puzzle posed by proclitics seems to be the following: though there is clear evidence that enclitics in EP are syntactically opaque and attached to the verb in the morphology (ch.4-7), proclitics seem to be syntactically visible like lexical units: there is, on the one hand, the interpolation of syntactic material (even if limited to Xº elements) (8.1.1) and, on the other, the ability of proclitics to have wide scope over coordinated VPs (8.1.2). Scopal and distributional properties of proclitics then clearly indicate that they are not morphologically part of the verb.

Given that proclitics cannot be attached to a verbal stem like enclitics, I shall now examine in more detail the properties of proclitics with the aim of determining their grammatical status.

8.2 Affixal evidence

The aim of this section will be to show that despite their phrasal behaviour, proclitics constitute affixal elements like enclitics. The first section examines formal similarities between enclitics and proclitics which indicate that proclitics form an affixal unit (8.2.1); the second provides evidence which weakens the view that proclitics constitute lexical items (8.2.2). In section 8.3 I shall then argue that while enclitics are morphologically attached to the verbal host, proclitics are best analysed as phrasal affixes.

8.2.1 Discussion

8.2.1.1 Clitic inventory

To start with, it is crucial to observe that the set of preverbal clitics coincides in form *and* content with the set of postverbal clitics. This means that the inventory of clitics presented in Chapter 6, represented here under (25), is representative of *both* enclitics and proclitics.

(25) Clitic inventory of EP

	REFLEXIVE	DATIVE	ACCUSATIVE
1.Sg.	me	me	me
2.Sg.	te	te	te
3.Sg.Masc.			o
3.Sg.Fem.	se	lhe	a
1.Pl.	nos	nos	nos
2.Pl.	vos	vos	vos
3.Pl.Masc			os
3.Pl.Fem.	se	lhes	as

To derive enclitics and proclitics as categorially different elements, it would be difficult to explain why the grammar would generate two complete sets of exactly identical units through completely distinct mechanisms (i.e., lexical entries, for proclitics, and realisation rules, for enclitics).

This point can be clarified with the case of Welsh where enclitics also behave distributionally and morphophonologically like affixes, whereas proclitics display a more syntactic-like behaviour. Sadler (1997) takes these differences to suggest that enclitics are affixes, and that proclitics are instead syntactic (non-projecting) units. For Welsh, however, the view that there are two sets of categorially distinct clitics seems to be conceptually well motivated, for enclitics and proclitics in this language are phonologically distinct from each other:

(26) Partial paradigm of Welsh clitic pronouns (Sadler 1997a)

	proclitics	Enclitics
1Sg	fy	‘m
2Sg	dy	’th
3Sg.Fem	ei	‘w, ‘I
3Sg.Masc	ei	‘w, ‘I
3Sg	-	‘s

For EP, on the contrary, treating proclitics and enclitics as categorially distinct units would completely fail to explain why there are absolutely no formal differences between the inventory of proclitics and the inventory of enclitics. For conceptual reasons, therefore, the EP enclitics and proclitics should be regarded as being categorially identical. I therefore assume that a unified analysis of preverbal and postverbal clitics should generate EP clitics as affixes, regardless of whether they appear preverbally or postverbally.

8.2.1.2 Clitic cluster

Further supporting the claim that proclitics should be analysed as affixes is also the fact that clitic clusters appearing in postverbal position are exactly like clusters in preverbal position. This homophony and structural parallel is illustrated briefly in (27) with portmanteau clusters (27a), rigid ordering (27b) and cluster-internal allomorphy (27c):

(27)

	postverbal position		preverbal position
a.	deu -mo	→	não mo deu
	gave -1sg.dat/3sg.masc.acc		not 1sg.dat/3sg.masc.acc gave
	‘(s/he) gave it to me’		‘(s/he) didn’t give it to me’
b.	entregou -se -lhe	→	até se lhe entregou
	deliver -3sg.ref -3sg.dat		even 3sg.ref 3sg.dat deliver
	‘it was delivered to him’		‘it has even been delivered to him’
c.	deste -no -lo	→	até no -lo deste
	gave 2sg.dat -3sg.masc.dat		even 2sg.dat -3sg.masc.dat gave
	‘(s/he) gave it to us’		‘(s/he) even gave it to us’

As argued in Chapter 4, cluster-internal properties such as portmanteau clusters, rigid ordering and cluster-internal allomorphy strongly suggests that clitic clusters should not be derived in the syntax and that they are best viewed as sequences of affixes (cf. Chapter 6). An inflectional analysis of cluster formation shall be provided in Chapter 9 which in fact assumes that cluster-internal regularities resemble regularities found in inflectional morphology.

On the contrary, a syntactic derivation of the proclitic clusters would be faced with the serious problem of accounting for these non-syntactic properties. As argued in previous chapters, the idiosyncratic structure of clitic clusters does not follow from general syntactic principles. In addition, if such clitic-specific principles were provided, then any syntactic derivation would also be faced with the difficult task of explaining why the grammar derives sequences of clitics – which are in all respects phonologically and structurally parallel – both in the morphology, for enclitic clusters, and in the syntax, for proclitic clusters.

We may therefore conclude that there are neither conceptual nor empirical reasons for treating proclitics and enclitics as categorially distinct units. To us, the fact that preverbal clusters are exactly identical to postverbal clusters clearly indicates that they constitute the same sequence of affixes.

8.2.1.3 Summary

Despite the the scopal and distributional behaviour of proclitics, proclitics are formally exactly identical to enclitics. This claim has been supported by the properties of the proclitic inventory (8.2.1.1) and the internal structure of the proclitic cluster (8.2.1.2) which suggest that both enclitics and proclitics are not categorially distinct. If this claim is correct, then both proclitics and enclitics must be generated as affixes in the morphology and regarded as positional variants of the same affixal unit.

So, we appear to have a number of points of similarity with enclitics (such as clitic homophony and cluster identity) which show that pronominal object clitics (regardless of whether they appear preverbally or postverbally) are in fact affixes. The question we need to address now is as follows: If

proclitics are generated in the morphology like enclitics, then how can we account for the syntactic behaviour of the proclitic-verb combination? Anticipating the conclusion somewhat, I shall argue that the 'mixed' properties of EP proclitics can only be adequately accounted for if proclitics are analysed as phrasal affixes. Within this analysis, the difference between enclitics and proclitics is not one of kind, but of domain of clitic placement. Adopting Klavans' (1980) and Anderson's (1992) distinction between word-level and phrase-level affixation, I shall assume that enclitics, as argued in Chapters 6 and 7, attach in the morphology as word-level suffixes, while proclitics attach syntactically to the left edge of a verbal domain. By assuming that proclitics select a phrasal node in the syntax, this hypothesis will be able to account for both the formal similarities between enclitics and proclitics and their distributional/scopal differences.

8.2.2 More evidence

This section further weakens the claim by Vigário (1999b) and Gerlach (2001a) that proclitics should be represented as independent syntactic nodes.

8.2.2.1 Data

a) Distribution

To start with, if proclitics were lexical units one would expect them to take part in syntactic operations generally associated with full pronouns, such as coordination, topicalisation or modification (Kayne 1975, Cardinaletti&Starke 1999). This prediction however is not confirmed, for EP proclitics seem to share with enclitics the fact that they cannot be coordinated (28a), topicalised (29a), modified (30a), or bear contrastive stress (31a):

(28) a. *O Pedro não *me* *e* *lhe* deu o livro
the Pedro not 1sg.dat and 3sg.dat gave the book
'Pedro didn't give me or him/her the book'

b. *Eu e ele* andámos na mesma escola.
I and he went to-the same school
'Me and him went to the same school'

(29) a. **Lhe,* o Pedro não deu o livro.
3sg.dat, the Pedro not gave the book
'Pedro didn't give the book to HIM'

b. *Dele*, sei pouco
of-him, know little
'I know very little about HIM'

(30) a. *Nós sabemos que *lhe* *apenas* deste um beijo.
we know that 3sg.dat only gave a kiss
'We know that you only gave her a kiss'

b. A prenda é para *ela* *apenas*.
the present is for she only
'The present is only for her'

(31) a. *Todos ME e não TE deram presentes.
all 1sg.dat and not 2sg.dat gave presents
'Everyone gave ME - but not YOU – presents'

b. Foi ELA e não ELE que partiu a janela
was she and not he who broke the window
'It was HER and not HIM who broke the window'.

The data then shows that the behaviour of proclitics does not follow from general syntactic principles. Similar criteria are used for enclitics in Chapter 3 to argue that clitics do not behave like full pronouns.

b) Interpolation revisited

Further supporting the idea that proclitics do not behave like lexical items is the restricted nature of interpolation. As argued before (cf. 8.1.1), interpolated elements cannot project into XPs and therefore the verb and the proclitic can only be separated by non-projecting Xº level elements, not by whole phrases. Indeed, even though interpolated elements such as *não* and the monosyllabic adverbials discussed in 8.1.1 behave like word-level units, there are severe restrictions on what these elements can do. For example, these units cannot coordinate, be modified or take complementisers (32a-34a), unlike full pronouns or adverbials occurring elsewhere in the clause (32b-34b)

(32) a. *Quantas vezes te *eles* *e* *elas* disseram que ....
how-many times 2sg.dat they.masc and they.fem said that ...
'How many times did they both tell you that ...'

b. *Eles* *e* *elas* fazem uma boa equipa

they.masc and they.fem make a good team.
'They make a good team'

(33) a. *Quantas vezes te *até eu* tenho de dizer para ...?
how-many times 2sg.dat even I have to tell that ...
'How many times do even I have to tell you to ...?'

b. *Até eu* te posso ajudar
even I 2sg.acc can help
'Even I can help you'

(34) a. *Quantas vezes te *nós, que gostamos de ti,* dissemos para comeres
how-many times 2sg.acc we who like of you, said to eat
'How many times did we, who like you, tell you to eat?'

b. Nós, que gostamos de ti, nunca te magoaremos.
we who like of you never 2sg.acc hurt
'We, who like you, will never hurt you.'

What this data then shows is that the relation between the verb and the proclitic is very local, and that proclitics, like enclitics, must gravitate around the verb. This high degree of locality between proclitics and the verb, thus indicates that the behaviour of proclitics is most similar to that of lexical items.

c) Wide scope affixation

In addition, the ability to be shared over coordinated phrases – though typical of words – should not be used as a bullet-proof criterion for word-status (cf. Miller 1992 for discussion). Wide-scope affixation is also attested cross-linguistically. The examples below show that case and number affixes can have wide scope over coordinated phrases and (semantically unrelated) coordinated words.

Inflectional morphology in Turkish seems to assign affixes the ability to be shared under coordination. For example, in (35), the affix cluster *-ʲer-im-i*, which is realising the exponents for plural, 1st singular possessive and accusative case, is shared by both conjuncts of the coordinated noun phrase *kʲedi ve kʲpøekʲ* 'cat and dog'. Likewise, in (36) the second plural agreement marker -

sunuz is shared by two verb phrases, namely *hastane-ye gid-iyor* 'going to the hospital' and *o-nu gör-üyor* 'seeing him/her'.

(35) [kʲedi ve kʲpøekʲ]-lʲer-im-i (Turkish)
cat and dog PL-1S.POSS-ACC
"my cats and dogs(acc)" (Orgun 1996)

(36) [hastane-ye gid-iyor, o-nu gör-üyor] -sunuz (Turkish)
[hospital-DAT go-PROG 3SG-ACC see-PROG] -2PL
"You all are going to the hospital and seeing him/her" {Yu&Good 2000}

Another example comes from Mari (Luutonen 1997) where morphological case markers, which generally occur on each coordinated noun, may also take wide scope over a conjoined noun phrase, as in (37b).

(37) a. puškuδê-lan δa joltaš-lan
neighbour-DAT and friend-DAT
"to the neighbour and friend"

b. puškuδo δa joltaš-lan
neighbour and friend-DAT

The case of EP proclitics then resembles that of wide-scope affixation. What does the data tell us about the grammatical status of proclitics? First, that wide scope does not rule out the affixal status of preverbal clitics. Second, that proclitics cannot be morphologically part of an adjacent host, for in that case it could not have wide scope.

8.2.2.2 Summary

The data surveyed in this section has shown that there is insufficient evidence in favour of viewing proclitics as autonomous syntactic words. I shall therefore explore the idea that proclitics constitute in effect 'phrasal prefixes' (Luís 2001a, 2003a, to appear; Luís&Sadler 2003). Under this view, the

asymmetry between enclitics and proclitics can be analysed as a difference in status between the preverbal (i.e. phrasal) and postverbal (i.e. word-level) placement of the same pronominal affixes.

Particularly revealing is the data in (38-39) which illustrates the behaviour of clitic clusters in clauses with interpolation and in coordinated structures. Assuming that clitic clusters constitute inflectional sequences, as argued so far in this thesis (following Miller&Sag 1997, Monachesi 1999, Auger 19994, Crysmann 1997, Spencer 2000, among others), then the fact that clusters exhibit phrasal behaviour further strengthens the claim that these affixes can also attach to a phrasal node. In (38-39 both interpolation and wide scope affect not only single occurrences of preverbal clitics (cf. 8.1) but also the portmanteau cluster *lho* which is separated from the verb by the negation marker 'não' in (38), and shared over two coordinated VPs in (39).

(38) Já lho não podemos dar.
anymore 3sg.dat-3sg.acc.masc not can give
'We cannot give it to him anymore'

(39) Nós sabemos que a Maria lho pediu de manhã
we know that the M. 3sg.dat/3sg.masc.acc requested in-the morning
e devolveu à noite.
and returned in-the evening
'We know that M. requested it from her/him in the morning and returned it to her in the evening.

Assuming that the cluster *lho* is derived either as one single affix or as two fused affixes, (cf. Chapter 7), the data in (35-36) clearly shows that even though proclitics are not attached morphologically to the verb, they must be generated as affixal elements.

8.3 Phrasal affixation

Phrasal affixes have been used in the literature as well-established inflectional categories (Klavans 1980, 1985, Anderson 1992, 1995, 2000, Spencer 2000, Legendre 2000a,b, Börjars 2003). In what

follows it is argued that the concept of phrasal affixation shall prove crucial in accounting for the similarities and differences between enclitics and proclitics. In particular, the claim that proclitics are *affixes* enables us explain why they are phonologically exactly identical to enclitics; and as *phrasal* affixes, we are able to accommodate the phrasal behaviour surveyed in 8.1.

Section 8.3.1 surveys the classical models of phrasal affixation by Klavans (1980) and Anderson (1992). Section 8.3.2 offers an intuitive proposal for the linearisation of proclitics. Section 8.3.3 then summarises previous accounts of EP proclisis and compares them to the proposal made in this thesis.

8.3.1 Models of phrasal affixation

As defined by Klavans (1980, 1985) and Anderson (1992), phrasal affixes are bound to syntactically defined positions rather than to word-level hosts (cf. Chapter 3) [52]. Therefore, both in Klavans (1980, 1985) and Anderson (1992), emphasis is given to the placement of phrasal affixes and to the ways in which placement can be determined[53].

Klavans (1980) provides one of the first accounts of the distribution of phrasal affixes. Based on the distributional properties of clitics cross-linguistically, she classifies clitics on the basis of the phrasal domain within which the clitic host occurs in combination with the three additional placement parameters. Parameter 1 determines whether the clitic attaches to the initial or the final constituent of the phrase which constitutes its domain; Parameter 2 specifies whether the clitic attaches before or after the host specified by Parameter 1, and Parameter 3 indicates whether the

---

[52] As mentioned in before in Chapter 3, the terms 'phrasal affixation' is not synonymous with 'edge inflection'. Edge inflections are morphologically part of the host, even though they tend to appear at the edge phrase (cf. English possessive, Bulgarian and Macedonian determiners in Halpern 1995).

[53] Cf. also Spencer 1992 for detailed summary of Klavans (1980, 1985) and Halpern (1998) for summary of Anderson (1992).

clitic will attach phonologically to the host to its left or to its right. The values for each parameter then are Initial/Final for P1, After/Before for P2, and Proclitic/Enclitic for P3.

Ngiyambaa clitics, as analysed by Klavans (1980), attach after the first word/constituent of the sentence. Klavans assumes that it takes the sentence as its domain and that the values for each one of the three parameters are Initial, After, Enclitic. Depending on how each parameter is set, the theory predicts eight types of clitics (see Spencer 1992 for detailed discussion of all eighth possibilities), although not all clitic types predicted by Klavans have been attested. This partly explains why alternative approaches to phrasal affixation have replaced Klavans' taxonomy. One interesting predictions made by Klavans taxonomy is the fact that P3 (expressing phonological liaison) can be independent of P1 and P2 in that the phonological host of a phrasal affix may not be contained in its syntactic domain. This is observed in the Australian language Nganhcara (Klavans 1980) where clitics occur before the final node under S, but attach phonologically as an enclitic. Such a mismatch between phrasal syntax and phonological host will be briefly illustrated with Cappadoccian and Romanian in Chapter 9).

Adopting some of Klavans' insights, Anderson (1992), develops a taxonomy for the classification of phrasal affixes making use of three parameters: scope, anchor and orientation. The 'scope' parameter expresses the domain within which the phrasal affix is located (e.g. S, VP, NP). The 'anchor' parameter defines the syntactic host with respect to which the phrasal affix is situated: it is assumed that it is positioned by reference to the FIRST, HEAD or LAST element within the phrasal domain. The third parameter defines the orientation of the phrasal affix; it expresses whether the phrasal affix PRECEDES or FOLLOWS the anchor point. Unlike Klavans (1980, 1985), the direction of phonological attachment is assumed to follow from inherent properties of the clitic and is not defined through an individual parameter as in Klavans. Illustrating briefly, Anderson's parameters might be applied as follows. In Kwakwala, case-marking clitics precede the noun they mark within the NP but are phonologically enclitic on the word preceding them. The placement

values for these phrasal affixes would be NP, Initial, Precedes. For Serbo-Croatian 2P clitics (cf. Chapter 3), the values would be VP, First, Follows (Halpern 1995).

The general idea that phrasal affixes attach to syntactic constituents has gained wide acceptance in linguistics among morphologists. However the fact that not all of the phrasal affix types predicted by Klavans (1980) and Anderson (1992) have been attested (cf. Spencer 1992, Halpern 2000), partly explains why current approaches to phrasal affixation have replaced the rigid parameters by language specific placement constraints. For example, the distribution of phrasal affixes has been more recently formulated in terms of Optimality Theoretic alignment constraints. Within Optimality Theory (McCarthy&Prince 1993a), edge alignment constraints are introduced which position the phrasal affix with respect to a phrasal host (e.g., Anderson 1995, 1996; Legendre 2000 and Spencer 2000). In Anderson (1996, in press), for example, clitic placement is derived from the ranking between Edgemost and Non-Initial constraints. So, whether clitics are enclitic or proclitic is determined by Edgemost (e,L) and Edgemost (e,R) where 'e' expresses the elements with respect to which the placement is determined. If Non-Initial (e) dominates Edgemost (e,L) then the clitic appears in second position, preventing it from appearing in clause-initial position. Anderson also provides constraints that are based on the integrity of constituents, such as Integrity (Word) and Integrity (XP), which help determine whether 2P clitics attach after the first word or the first phrase. (cf. Legendre 2000 for alternative OT proposal).

As the above summary has shown, regardless of how phrasal affixes become ultimately situated, their position must be defined in terms of a phrasal domain and a syntactic anchor point inside that domain. In purely morphological terms, this means that phrasal affixes do not form a morphologically cohering unit with an adjacent free word, unlike ordinary morphological affixes (cf Chapter 9 for discussion and inflectional account).

For EP, however, the claim that both enclitics and proclitics should attach to a phrasal node, as suggested by Anderson (2000), raises obvious empirical problems. As previous chapters of this

thesis have shown, enclitics cannot be separated form the verb, they trigger stem allomorphy on the verb, undergo allomorphic variation[54], attach to the infinitival stem before the tense/ agreement endings, and cannot have wide scope. These properties indicate that EP enclitics shosuld be regarded as genuine suffixes. Note that phrasal affixes should not trigger allomorphic effects given that they are not attached in the morphology and do not form a morphologically cohering unit with the host. In addition, phrasal affixes can typically have wide scope and they do not select the category of their host[55]. Assuming the basic difference between phrasal affixes, on the one hand, and morphological affixes, on the other, I take the view that a uniform analysis of EP cliticsation based purely on phrasal affixation is not tenable (cf. Chapter 3). I therefore conclude that only EP proclitics should be viewed as phrasal affixes.

8.3.2 EP proclitics

The models of phrasal affixation surveyed above have shown that the typology of phrasal affixes depends on the distributional properties they exhibit. This section will briefly examine the linearisation of EP proclitics and argue that proclitics are V-oriented phrasal affixes.

8.3.2.1 Non-initial position

One of the puzzles of EP proclitics is the fact that they never occur in clause-initial position, unlike Spanish or Italian (cf. Chapter 10 on proclitic contexts). To account for this fact, previous syntactic treatment of clitic pronouns have suggested that clitics in EP are attracted to the 2P of the clause

---

[54] In fact, the ability of clitics to trigger allomorphy has also been used by Halpern (1995) to support the view that Macedonian and Bulgarian determiners are morphologically part of the host. Similar arguments have been adduced for the English Possessive marker (cf. footnote 1).

[55] Further problems with Anderson's analysis include the fact that clitic placement is assumed to be triggered by the Tobler-Mussafia Law. Problems with this assumption are addressed in Chapter 10.

(Rouveret 1992, Spencer 1992, Madeira 1993, Salvi 1990). If we look the data in (40), we effectively observe that proclitics appear in second position, that is, "after the first word", under the original definition of Wackernagel (1892).

(40) a. Não *lhes* mostraram os presentes.
not 3pl.dat show the presents
"They didn't show them their presents."

b. Até *o* viram em casa.
even 1sg.masc.acc saw at home
"They even saw him at home."

c. Quantos *nos* compraste?
how-many 2pl.dat show
"How many did you buy us?"

d. Dele *se* sabe pouco.
of-him 3sg.refl know little
"Little is known of him."

e. Todos *as* esconderam meticulosamente.
all 3pl.fem.acc hide meticulously
"They all hid them meticulously."

f. Que *se* vendem bem, sabe ele há muito tempo.
that 3pl.refl sell well, knows he for much time
"That they sell well, is something he has known for a long time."

Even though a 2P analysis would explain why proclitics are never first in the clause, as the data in (40) illustrates, it is far from clear that the Wackernagel Law would accurately encapsulate their distribution. First of all, there is no restriction forcing the proclitics to appear in 2P rather than further away from the edge of the clause. As such, appearing in 2P constitutes only one of the many preverbal positions proclitics may occupy. For example, as shown in (41), proclitics may also be third or fourth in the clause.

(41) a. [As crianças , por vezes, não] *lhes* mostraram os presentes.
the children for times not dat.3pl showed the presents
"The children sometimes didn't show them their presents."

b. [As crianças até] *o* viram em casa.
the children even acc.1sg.masc saw at home
"The children even saw him at home."

c. [Dele ainda] não *se* sabe muito.
of-him yet not refl.3sg know little
"Still very little is known of him"

d. [Todos os alunos imediatamente] *as* esconderam .
all the students, of immediately, acc.3pl.fem hide
"All the students immediately hid them away meticulously."

e. ... [que os livros] *se* vendem bem.
... that the books refl.3sg sell well
"that the books sell well."

There is in effect no clear restriction on the distance that can exist between the edge and the proclitic position. As in (42), the proclitic can be moved further away from second position if we coordinate, specify or modify the subject.

(42) a. [Os livros que eu li não] *me* agradaram.
The books that I read not 2s.dat pleased
'The books I read didn't please me'

b. [As tuas duas filhas e os meus dois filhos até] *lhe* agradeceram.
the your two daughters and the mine two sons even 3sg.dat thanked
'Your two daughters and my two sons even thanked him/her'

It seems that the limit is solely imposed by the verb, which is preceded by the proclitic.

It may, of course, be argued that the Wackernagel position in EP should be defined in terms of constituents, rather than in terms of single words. In that case, 2P could mean "after the fist phrase". However, even if the 'law' is rephrased in terms of constituent structure, we would still not be able

to find clear cut data supporting the claim. For example, while in typical 2P languages the complementiser is immediately followed by the clitic (Klavans 1980, Anderson 1992, Spencer 1992), this requirement is only optional for EP. It might also be argued that proclitics are only residually 2P and that any deviations from canonical 2P placement follow from other regularities or simply from ad-hoc stipulations. However, in the absence of straightforward evidence, I shall assume that the inability of proclitics to appear sentence-initially cannot be regarded as a Wackernagel effect.

8.3.2.2 V-oriented placement

This section argues that EP proclitics should instead be classified as V-oriented phrasal affixes That is, phrasal affixes which select a verbal anchor point. The insight that proclitics are attracted to the verb is also adopted by Duarte&Matos (2000) and Crysmann (1999). As V-oriented phrasal affixes, I assume that there is a high degree of locality between proclitics and the verb. Supporting this view is the fact that proclitics must always appear in the vicinity of the verb. They can only be separated from the verb by single interpolated words (43), never by phrasal constituents (44b). The latter can also appear before the verb, but they must precede the proclitic (44a).

(43) a. [que] *o* ajudou
[that] 3sg.masc.acc helped
'[that] (s/he) helped him'

b. [que] *me* não ajudou.
[that] 1sg.dat not helped
'[that] s/he didn't help me'

c. Quantas vezes *lho* não dissemos?
how-many times 3sg.dat/3sg.masc.acc not told
'How many times didn't we tell it to him?'

(44) a. Sei que [ainda na semana passada] *te* telefonou.
(I)know that [already in-the last week] 2sg.acc phoned
"I know that s/he already phoned you last week."

b. *Sei que *te* [ainda na semana passada] telefonou.
(I)know that 2sg.acc [already in-the last week] phoned

Let us therefore assume that the phrasal domain of the proclitic will be the VP domain. Within this VP domain, we need to determine the anchor point of the proclitic. As mentioned before, proclitics may take scope over coordinated VP, but it may also be repeated on each conjunct like enclitics. I shall extend the two-way distinction between X^{max} and Xº clitics proposed by Halpern&Fontana (1994) to phrasal affixes and assume that proclitics attach to the left of Vº when they have narrow scope and to the left of VP when they have wide scope[56]. The anchor point for EP proclitics shall therefore be defined as a V-VP pair. Under this assumption, we also capture the high degree of locality between the verb and the proclitic, both for proclitics with narrow scope and wide scope. In either case, we assume that Adverbial Phrases adjoin to VP, leaving the verb in VP-initial position. Evidence supporting the adjunction of ADvP comes from coordinated VPs which share the same AdvP.

Determining the exact constituents depends on the theory of syntax one adopts and on the assumptions one makes about the structure of EP clauses. For concreteness, I assume that the EP verb is under V within VP, however this is in conflict with the usual assumption within the

[56] Some of the scopal effects found with proclitic constructions are dicussed in Luís&Sadler (2003) within Lexical-Functional Grammar. However how the precise nature of the constraints on coordinated clitic structures remains to be fully investigated. Other accounts within different frameworks, include Matos (2000) within derivational syntax and Crysmann (2002) within HPSG.

Principles and Parameters/Minimalist literature that finite verbs occupy I, not V (Costa 2000, Matos 2000). Within this phrase-structure representation of the EP clause, the obvious choice for proclitics would be I´. I will however remain agnostic as to whether the verb is in I or in V.

Regardless of which configuration is adopted, the crucial aspect about proclisis is that phrasal affixes in preverbal position do not form a morphologically cohering unit with the host they attach to. The host is a phrasal node and not a morphological stem. This view is also shared by Spencer (2000), Legendre (2000a,b), Börjars (2003). Unlike proclitics, EP enclitics constitute morphological suffixes and form with the host an inflected verb form (cf. Chapter 6). This insight is also assumed for pronominal clitics in French (Auger 1994), Macedonian (Spencer 2000), Spanish (Andrews 1990) and for incorporated pronouns in Chichewa (Bresnan&Mchombo 1987).

As to the restriction against clause-initial placement, I will assume with Martins (1994), Duarte&Matos (2000), Crysmann (1997, 1999) that proclitics must be preceded by any one of the proclitic triggers given in (40-41). The preverbal elements which prevent proclitics from appearing as the first element of the clause include clausal negation (cf. 40a), certain aspectual adverbs (40b), wh-phrases (40c), focused phrases (40d), some quantifiers (40e), and sentential complementisers (40f). We shall return to this data in Chapter 10 and argue that the syntactic conditioning of proclisis can be acccommodated within the inflectional analysis of clitics provided in this thesis.

8.4 Summary

This chapter examined the behaviour of proclitics and argued that, they share with enclitics a wide range of properties which suggests that they are best generated as affixes. Based on the morphological properties, I argued that proclitic clusters and enclitic clusters correspond to the same sequence of affixal clitics.

However, their behaviour with respect to interpolation (8.1.1) and scope (8.1.2) indicate that, unlike enclitics, proclitics cannot form a morphologically cohering unit with the verb. To capture the

asymmetry between clitics and proclitics, it has been argued that proclitics (and proclitic clusters) are really phrasal prefixes: affixes with special linearisation properties. In particular, they linearise in the syntax with respect to a phrasal domain; in the literature, the claim that affixes in some languages select phrasal hosts has been proposed for a wide range of languages (Klavans 1980, Anderson 1992, Legendre 2000, Spencer 2000). Enclitics in EP, on the contrary, constitute genuine verbal suffixes (i.e. in terms of placement, they attach to verbal stems). Unlike the EP clitic system, clitic pronouns in other Romance languages behave uniformly like morphologically cohering affixes (Miller&Sag 1997, Monachesi 1999, Auger 1994, among others). An analysis capturing these insights will be developed in the next chapter within PFM. The proposal assigns a given affixal clitic the ability to be placed either as a phrasal affix, for proclitics, or as a morphological affix, for enclitics.

Chapter 9 Clitic clusters within Paradigm-Function Morphology

Regardless of where in grammar affixal clitics attach, either in the morphology or in the syntax, we need to be able to explain, first of all, why affixal clitics can appear before or after a given host, as proclitics and enclitics. At the centre of this chapter, therefore, is the dual placement of affixes, a property which is not exclusive to clitic systems but has also been found in ordinary inflectional systems (Stump 1993, Noyer 1994). In our analysis of placement, it will become clear that placement cannot be dissociated from cluster formation, specially because clitic order inside the cluster is not affected by dual placement, as (1), illustrates. Thus any analysis of clitic placement must ensure that the unit which is positioned in enclitic and proclitic position is exactly identical.

(1) a. apresentou *-se* *-lhe*
introduced -3sg.refl -3sg.dat
'introduced himself/herself to him/her'

b. (não) *se* *lhe* apresentou
not 3sg.refl 3sg.dat introduced
'didn't introduced himself/herself to him/her'

The nature of clitic ordering will lead us to conclude that the one-by-one realisation of affixes, as standardly adopted within Paradigm-Function Morphology (Stump 2001), fails to correctly derive clitic sequences. An alternative analysis of clusters formation therefore shall be proposed which derives clitic clusters as whole units, following insights from previous studies on this topic. To account for dual placement of clitic clusters, I shall adopt insights from Stump (1993), on ambifixal position classes and argue that the direction of attachment of a given affix should be determined by independent placement rules, rather than by the realisation rules generating affixes. It shall be

further argued that placement rules apply only once to the whole cluster rather than to each individual clitic, in order to capture the fact that one and the same sequence of affixal clitics can appear before or after a given host.

The chapter proceeds as follows: section 9.1 addresses the dual placement of affixal clitics in Romance and EP. A preliminary analysis is formulated which explores Stump's (1993) insight about the 'separation' between realisation and placement. Section 9.2 then extends the discussion to clitic sequences and argues that clusters must be generated as a whole prior, through function composition (Spencer 2000). Section 9.3 then combines the 'separation' hypothesis with function composition: clitic clusters are derived as composed inflectional units, and placement rules are assumed to apply only once to the whole cluster. This account nicely captures the asymmetric nature of clitic placement discussed in the previous chapter, because it shall be assumed that the alignment function placing clitic clusters defines the domain within which clitics attach. In 9.4, cross-linguistic evidence suggests that asymmetric placement is not unique to EP. The analysis is captured within a modified version of the theory of Paradigm-Function Morphology, based originally on Spencer (2000) and further developed by Luís (2001a, 2002b, 2003b), Luís&Spencer (in press) and Spencer (ms).

9.1 Dual placement

As noted by Halpern (1995), one can investigate clitic clusters from the point of view of their internal or external distribution. The external distribution refers to the relation between clitics and the host, while the internal distribution focuses on how clitics relate to each other inside the cluster. As this section will show, it is somewhat straightforward to derive the external behaviour of clitics from the point of view of the theory of Paradigm-Function Morphology (Stump 2001). However when we attempt to combine the external with the internal behaviour, as shall be done in section 9.2, we realise that standard assumptions about affix realisation are not adequate.

Before considering the internal distribution of clitics, this section will address the dual placement of clitics, i.e., the ability for one and the same affixal unit to be positioned before and after the host (9.1.1). I will also show that Stump (1993) offers a conceptual basis for capturing dual placement (9.1.2) and that with minor modifications it is also possible to capture asymmetric affixation, i.e., phrasal and morphological attachment of the same affixal unit (9.1.3).

9.1.1 Preliminaries

To start with, we shall return to some of the assumptions underlying the analysis of suffixing clitics presented in Chapter 6. A fragment of the realisation rules adopted for the derivation of clitics is given in (2) for each one of the three rule blocks.

(2) a. $RR_{A1,\ \{Case:\ Acc,\ Refl:+,\ Person:3\},V}$ (<X,σ>) $=_{def}$ <X-se>

b. $RR_{B5,\ \{Case:\ Dat,\ Person:3,\ Number:\ Sg\},V}$ (<X,σ>) $=_{def}$ <X-lhe>

c. $RR_{C1,\ \{Case:\ Acc,\ Person:3,\ Number:\ Sg,\ Gender:\ masc\},V}$ (<X,σ>) $=_{def}$ <X-o>

Based on the order of application of these rules, it is assumed that a postverbal cluster such as *se-lhe* is derived as illustrated in (3) where X in (3a) stands for an inflected verb (previously derived as discussed in Chapter 6) to which RR_{A1} from rule block A applies; the output of applying RR_{A1} to X serves as input to RR_{B5} from block B as in (3b): RR_{B5} applies to X-*se* yielding X-se-*lhe*. Finally, since rule block C provides accusative clitics, it shall not apply in the derivation of *se-lhe*; therefore, by the Identity Function Default block C applies vacuously (see Chapter 6).

(3) partial derivation of enclitic cluster *se-lhe* (cf. Chapter 6)

	input	output	
a.	X	X-se	(by RR_{A1})
b.	X-se	X-se-lhe	(by RR_{B5})
c.	X-se-lhe	X-se-lhe	(vacuous)

Essentially what we see is that ordinary RRs, such as those proposed in (2), apply to the output of the previous RR. This is illustrated in (4) where each RR is defined over a morphological base which constitutes the anchor point to which inflectional affixes attach. The anchor can be a root or a stem, or simply an underived base (i.e., the output of the previous RR):

(4) a. [stem]-x
b. [stem-x]-y
c. [stem-x-y]-z
d. etc.

Through the one-by-one realisation of suffixes, cliticisation in Chapter 6 has been derived in a sequential way. The analysis however constitutes only a partial approach to cliticisation, given that it does not contemplate the fact that clitics can also appear preverbally.

Let us therefore consider how preverbal clitics might be derived. As is well-known, the ability for one and the same affix to appear before and after the host is typical of Romance languages (perhaps with the exception of some informal varieties of Brazilian Portuguese). Based on the proposal in Chapter 6, it might be suggested that deriving preverbal counterparts is a straightforward matter. One simple way of obtaining a proclitic sequence might be to propose an additional sets of RRs, perhaps as in (5), which generate clitic affixes appearing before the host (the host is represented as X, but its exact identity is for the present discussion not relevant; see below for discussion about domain of placement).

(5) Fragment of RRs for proclitics (tentative)

a. $RR_{A1,\ \{Case:\ Acc,\ Refl:+,\ Person:3\},V}$ (<X,σ>) $=_{def}$ <se-X>
b. $RR_{B5,\ \{Case:\ Dat,\ Person:3,\ Number:\ Sg\},V}$ (<X,σ>) $=_{def}$ <lhe-X>
c. $RR_{C1,\ \{Case:\ Acc,\ Person:3,\ Number:\ Sg,\ Gender:\ masc\},V}$ (<X,σ>) $=_{def}$ <o-X>

Even though it appears to be possible to simply propose the sets of rules in (2) and (5), there are various conceptual and empirical reasons for not adopting this proposal. I shall only refer to one of the problems, and postpone the other for later. What seems to be missing from the account skeched in (2) and (5) is the insight that the set of enclitic suffixes is exactly identical to the set of proclitic prefixes. For French, deriving proclitics by 'duplicating' the RRs, as suggested in (5), would perhaps be appropriate because in this language proclitics can be phonologically distinct from enclitics. In French, for example, there are distinct forms for 1st and 2nd singular clitic pronouns depending on whether they occur preverbally or postverbally (Miller& Sag 1997). However, for the majority of the clitic pronouns of French and for the whole system of EP, the set of preverbal and postverbal affixes is completely homophonous, the number of proclitics is exactly identical to the number of enclitics, and both enclitics and proclitics convey the same morphosyntactic feature set. There is then insufficient evidence for proposing two distinct sets of rules - one for the prefix realisation and another for the suffix realisation. So, the first question we need to address in our treatment of clitic clusters is how to account for the fact that one and the same affix can appear before or after the host.

9.1.2 Dual affixes

Affixes which have the ability to be attached either as prefixes or suffixes are not only attested in clitic system, but have also been found in ordinary inflectional systems (e.g., Swahili and Huave) as shown by Stump (1993) and Noyer (1994). Within the theory of Paradigm-Function Morphology, Stump (1993) has defined this type of affix as ambifixal exponents, namely prefix/suffix pairs whose members are identical in phonological form and encode the same feature specifications (Stump 1992). This section summarises one of the treatments offered in Stump (1993) for handling ambifixal pairs. It is based on the assumption that the similarity between prefixal and suffixal exponents should be captured by separating the realisation of affixes from their placement. Under

this view, RRs do not specify whether the affix attaches as a prefix or as a suffix; independent placement rules that are associated with these RRs determine the placement of the affix[57]. How this proposal works for Swahili will now be illustrated. We shall then examine in 9.1.3 whether these proposals can be extended to cliticisation.

Swahili relative affixes agree in gender and number with the relativised argument, appearing before or after the verb depending on the tense and polarity features of the verb form. If the verb form is either marked for tense or negative polarity, it appears prefixally (6), but if the verb form is both tenseless and positive, it appears suffixally (7). Crucially, the group of prefixes is phonologically exactly identical to the group of suffixes and encodes the same feature specifications (cf. Stump 1993). Any account of mobile affixes should therefore capture that both groups of affixes are related, rather than deriving them as completely independent sets of affixes.

(6) mtu a-na-*ye*-soma (Swahili, Stump:139)
person SU.AGR-TNS-REL-read
'a person who is reading'

(7) mtu a-soma-*ye*
person SU.AGR-read-REL
'a person who reads'

To account for this type of affixal behaviour Stump suggests a different type of realisation rule, namely realisation rules which simply realise the affix as an exponent (phonological form) but which do not determine whether it is placed as a prefix or as a suffix; affix placemetn would be

[57] In Stump (1993), one alternative proposal is proposed to handle ambifixal position classes. It makes use of metarules (i.e., rules of referral) which determine that for every realisation rule deriving a suffix there is a realisation rule realising a prefix belonging to the same rule block. This means that the prefixal set of rules would be derived from the suffixal set. Again, the analysis attempts to capture the relation between ambifixal pairs without introducing two separate sets of RRs.

handled by independent placement rules. Illustrating this insight, the RR in (8) defines the phonological form of the Swahili affix *ye*. It is realsied as an exponent which has not been linearised with respect to its base, as the comma between *ye* and the stem indicates.

(8) Realisation rule for the relative affix (Swahili, Stump:174)

$RR_{II,\ \{Rel:sub/obj:\alpha,\ Num:\ Sg:\ Gender:\ m/wa\},V}\ (<X,\sigma>) =_{def} \{X,ye\}$

(9) Placement rules

a. prefixation: *ye* is placed before X, by default

b. suffixation: *ye* is placed after X, if {Tns:none, Pol:+, Rel:subj/obj, Num:β, Gen: α}

The placement rule in (9) specifies the placement of the affix, determing whether the affix appears before or after the verb (i.e., whether it is realised as a suffix or as a prefix). Note that the prefixation rule is associated with the more general set of features, while the suffixation rule is more narrowly defined. Under the Elsewhere Condition, the relative suffix will be spelled out in positive tenseless forms, whereas the prefix will be realised as the default. This view then is very different from the traditional format of RRs which conflates form and direction of placement. I will now examine how to extend Stump's proposal to affixal clitics in Romance and EP.

9.1.3 Enclisis and proclisis

This section provides a preliminary analysis of the enclisis-proclisis placement of affixal clitics. The separation-approach does indeed provide an account which explains why clitics and proclitics are exactly identical. In addition, with minimal modification it also enables us to determine the nature of the host to which affixal clitics attach. Later in this chapter, more precisely in section 9.2, it will be argue that further extensions need to be introduced in order to account for the rigid ordering of clitics.

a) Dual placement of affixal clitics

Under the separation-account, it would be possible to assume that there is only one RR generating the affixal clitic as illustrated in (10).

(10) Realisation rule for the clitic affix

$RR_{B5,\ \{Case:\ Dat,\ Person:3,\ Number:\ Sg\},V}\ (<X,\sigma>) \quad =_{def}\ \{X,\ lhe\}$

What (10) says is that *lhe* constitutes an inflectional exponent realising a given set of morphological features. This affix is not generated as being inherently suffixal or prefixal because it is not linearised with respect to the stem.

To determine the placement of this affix we would need to postulate the placement rules in (11), which position the affix preverbally in (11a) and postverbally in (11b). We will not address the exact conditions under which proclisis and enclisis takes place at present (see below); it suffixes to say that proclisis is more marked than enclisis.

(11) Placement rules (ignoring the factors determining placement)

a. proclisis: *lhe* is placed before X

b. enclisis: *lhe* is placed after X

The crucial aspect of such an analysis is that it derives the homophony between preverbal and postverbal clitics naturally. In addition, it also captures the fact that the affixal clitics realising the feature content in (10) are phonologically exactly identical.

b) Mixed attachment

In Chapter 8, it is argued that affixal clitics can attach either to a verbal stem or to a phrasal node. This property of affixes is not contemplated in Stump (1993). I will now show that an analysis along the line just sketched can also elegantly capture the ability for a given affix to attach either to a morphological host or to a phrasal host

Note that (11) makes very little assumptions about the nature of the host. The host is simply annotated as X, without any claims being made as to the domain or the type of entity clitics attach to. However we know that for the EP clitic system, the attachment is asymmetric given that enclitics select a verbal stem while proclitics attach to a Vº node. To capture this 'mixed' attachment we may define that X in (11a) refers to a phrasal base and that X in (11b) refers to a morphological stem, as shown provisionally in (12).

(12) Placement rules (ignoring the factors determining placement)
a. proclisis: *lhe* is placed before Vº-VP
b. enclisis: *lhe* is placed after Vstem

So, taking the assumption that the identity of the clitic host can be defined on a language- specific basis, we easily capture that one and the same clitic cluster can attach to either a morphological or a syntactic host. For most of the other Romance languages, in which affixal clitics attach uniformly to an inflected verb form (Monachesi 1999, Nishida 1987, Brines 2000, Miller&Sag 1997), the host would be defined accordingly.

So far then it has been argued that in order to capture the close resemblance between proclitic and enclitic affixes preverbal and postverbal clitics could be derived either through the separation between affix-realisation and affix-placement. This approach seems to be able to capture the insight that proclitics and enclitics constitute one and the same affixal unit. It also enables us to define the clitic host on a language-specific basis, accounting for the fact that languages with similar clitic

clusters, such as EP and the other Romance languages, can differ with respect to the grammatical domain within which clitics attach.

Note, however, that we have so far only addressed single occurrences of clitics. To determine whether any of these approaches can be adopted for cliticisation we first need to examine whether they capture the dual placement of *clitic sequences*.

9.2. Clitic clusters

9.2.1 Invariant clitic order

This section shows that the separation-approach, as proposed in the previous section, fails to capture the fact that clitic order inside the cluster remains invariant regardless of whether the cluster appears preverbally or postverbally. In what follows, I will illustrate the problems posed for clitic clusters.

A fragment of the RRs is given in (13-15) for each one of the clitic positions discussed in Chapter 6: block A provides 3rd person reflexive clitics and block B provides, among other clitics, 3rd person datives (for simplicity I leave out block C which provides all 3rd person accusative clitics). Block A applies before block B (and block B before block C).

(13) a. $RR_{A1,\ \{Case:\ Acc,\ Refl:+,\ Person:3\},V}\ (X,\sigma) =_{def} \langle X,se\rangle$

(14) a. $RR_{B5,\ \{Case:\ Dat,\ Person:3,\ Number:\ Sg\},V}\ (X,\sigma) =_{def} \langle X,lhe\rangle$

(15) a. $RR_{C1,\ \{Case:\ Acc,\ Person:3,\ Number:\ Sg,\ Gender:\ masc\},V}\ (X,\sigma) =_{def} \langle X,o\rangle$

We also assume that there are placement rules as suggested in (12) evaluating the position of the clitic. As shown in (16), the derivation of the postverbal clusters is unproblematic given the RR in (13-15) and the placement functions in (12): block A applies before block B, we derive the suffixing sequence *se-lhe* as in (16). Each RR is further defined by the placement function in (12b) and attached one-by-one as a suffix.

(16) Enclitic cluster (informal illustration)

a. X-se (by 15, 12b)

b. X-se-lhe (by 16, 12b)

c. X-se-lhe (vacuous realisation)

As will be shown next, the set of rules in (13-15) combined with (12) fail to correctly derive the preverbal sequence.

a) Hypothesis 1

The problem with the rules in (13-15) is that they replicate the rule block order of the postverbal rule blocks. This means that the same precedence relation is found regardless of whether the affixes attach as suffixes or as prefixes: in either case, reflexive clitics from block A precede dative clitics from block B and so forth. In other words, the clitic nearest to the host in postverbal position will also be realised as the nearest affix to the host in preverbal position, given that rule block order remains unchanged. This order however fails to derive the correct preverbal cluster as the incorrect sequence *lhe-se*-X in (17) illustrates.

(17) Proclitic cluster (informal illustration)

a. se-X (by 15, 12a)

b. lhe-se-X (by 16, 12a)

c. lhe-se-X (vacuous realisation)

Upon closer observation, we observe that the order in which rule blocks apply can only derive proclitic sequences as mirror images of enclitic sequences, thus failing to capture the fact that the order is in effect invariant as in (4) above.

To clearly appreciate the significance of the intriguing behaviour of clitic clusters, we can briefly compare how ordinary affixes would behave in similar circumstances. To do that, however, we need to imagine an ordinary inflectional system in which two formatives exhibiting dual placement,

say Aff1 and Aff2, can appear either all before or after a given stem. Assuming that Aff1 is realized by rule block 1 which applies before block 2, then we will expect that Aff1 is realised before Aff2 (cf. Chapter 2 and Chapter 6 on rule block ordering). This affix order is illustrated in (18a) for the suffixing sequence; and it is also illustrated in (18b) for the prefixal sequence; crucially, Aff1 is realised as the affix closer to the stem in both (18a) and (18b); and consequently, Aff2 is realized as the affix which follows the innermost one.

(18) a. stem-Aff1.Aff2
b. Aff2.Aff1-stem

b) Hypothesis 2

An alternative analysis might be possible if it is assumed that the order within which rules apply must be inverted.

It is not clear how such an analysis might be developed, but for the present discussion let us just imagine an account in which enclitics and proclitics applied in the opposite order. This would entail that rules from block A would apply before rules from block B if the clitics attached postverbally, and the inverse order would hold if clitics attached preverbally. Leaving aside the exact technical implications of this inversion, we would simply assume that the last rule to apply in postverbal position would be the first one to apply in preverbal position. This assumption would correctly capture the fact that while the reflexive clitic is closer to the host postverally it is further away from it preverbally (cf. (1) above). Interestingly however, evidence suggests that inverting rule block order does not constitute an alternative. Most of all, because it would completely fail to explain important cluster internal factors such as phonologically conditioned clitic allomorphy.

As mentioned to in Chapter 4, 3rd person accusative pronouns exhibit three allomorphs: the *n*-form, the *l*-form and the default form which is vowel-initial. The *l*-form is induced by a consonant-final verb (i.e., -s, -z, -r, except if the verb has a nasal diphthong) resulting in the deletion of the verb-final consonant (cf. 19). What is interesting is that this type of alternation, which can only be

induced by clitics, also occurs cluster-internally. In this case, the triggering consonant is that of the preceding clitic as exemplified in (20), where consonant-final clitic *nos* triggers the *l*-form of the accusative clitic. We can assume that the allomorphy found inside the cluster and also at the verb-enclitic boundary is best described as a progressive type of assimilation where the preceding context has an effect on the following segments. In other words, the consonant-final clitic determines the allomorph of the accusative clitic.

(19) levamo -lo (*levamos-o)
took -3sg.masc.acc
'we took it'

(20) O Paulo deu *-no* *-lo*. (*nos-o)
the P. gave -2.pl.dat -3.sg.masc.acc
'Paulo gave it to us'

Intriguingly, the alternation found in (20) also occurs preverbally as shown in (21).

(21) O Paulo não *no* *-lo* deu. (*nos-o)
the P. not 2.pl.dar -3.sg.masc.acc gave
'Paulo didn't give it to us'

This shape alternation seems to suggest that the cluster cannot be derived as a sequence of prefixes to avoid the realisation of the 3rd person accusative clitic *before* the dative clitic. The data then shows that allomorphic effects can only be explained if proclitics are derived as suffixal sequences.

One could argue that the allomorphy in preverbal position is feature-based and that what triggers the effects is the co-occurrence of 1st and 2nd plural pronouns with 3rd person accusatives, rather than as a phonologically conditioned effect. In realisation-based models of inflection, where the affixes are triggered by the overall feature context associated with the verb form, this data would not pose any problems. The problem with that, however, is that two different accounts would be given for the same type of shape alternation: prefixal clitics would require an explanation based on grammatically

conditioned allomorphy, while suffixal clitics would capture the shape variation as phonologically conditioned allomorphy.

Another argument can be adduced against inverting order of rule application. If affixes in proclitic position are positioned in the opposite order, then we would need to formulate co-occurrence restrictions in the opposite order, too. So, while we have so far argued that reflexives must precede datives and that datives precede 3rd person accusatives, for proclitic sequences one would have to say that datives precede reflexives. Again, the problems with this type of derivation is that we are using two different sets of co-occurrence restrictions for the description of what seems to be the same cluster unit[58].

Summing up, the data has shown that the one by one attachment of affixes derives proclitic clusters as mirror images of enclitic clusters. Since only one rule block order would be available for the realisation of both preverbal and postverbal affixes, the clitic closer to the host in postverbal position would also end up being closer to the host in proclitic position. We also examined whether rule block order might be reversed so as to avoid deriving 'mirror images', but this approach would miss important generalisations about allomorphic alternations and co-occurrence restrictions.

It seems that alternative ways of deriving clitic sequences need to be sought. In the next section, I survey inflectional approaches to clitic ordering which adopt quite different assumptions about clitic cluster realisation. Rather than introducing clitics on a one-by-one basis, clusters are treated as whole units.

[58] The same holds for any account of proclitic clusters based on the 'more' ordinary RRs suggested in section 9.1 (cf.5).

9.2.2 Earlier inflectional accounts

In this section, I shall consider two inflectional views of cluster formation. The first approach, by Monachesi (1999), is cast within HPSG and regards clusters as opaque portmanteau forms; the second, by Anderson (1996, in press), views clusters as transparent sequences of affixes. What is interesting is that, in both proposals, cluster formation is independent of cluster placement.

MONACHESI (1999)

Working within the theory of Head-Driven Phrase-Structure Grammar (Pollard&Sag 1992), Monachesi suggests that the featural combination of two clitics should be mapped onto a portmanteau affix, as illustrated below.

(22) Monachesi (1999)

$$\left[\begin{array}{l}\text{complex-morph} \\ \text{STEM| SS| L| CLTS <NP [dat] 1sg, NP [acc] 3sgm>} \rightarrow\end{array}\right] \left[\begin{array}{l}\text{affix} \\ \text{AFFIX} \left[\text{PHON|SKEL < melo>}\right]\end{array}\right]$$

Note that the cluster *melo* is not generated as an inherently suffixal or prefixal unit, but simply as an affix. It is assumed that this cluster can be placed either as a proclitic or as an enclitic sequence. This analysis then seems to solve the problem of rigid ordering by assuming that one and the same clitic sequences can be positioned preverbally or postverbally. We shall see that Anderson (in press) and Spencer (2000), among others, make similar assumptions about cluster formation to capture the correct order of clitics inside the cluster.

The problem with Monachesi's analysis however is that clusters are derived as opaque (i.e., non-analysable) forms. This is a view which is adopted by non-inflectional studies, e.g., Schwarze (2001) and Vigário (1999b), and which has also been proposed by Monachesi (1999) for Italian clitic pronouns. It is surprising to find this claim within an inflectional analysis, given that it is generally argued that most of the morphological effects found inside clitic clusters would follow

naturally form an affixal treatment of clitic pronouns (Miller&Sag 1997, Crysmann 1997, Spencer 2000, Luís 2001a, 2001b). Monachesi's claim that clusters are internally 'opaque' solves quite clearly the ordering paradox discussed in the previous section. If clusters have no internal structure, then the question of how to derive the clitic sequence sequentially is no longer relevant. However, the analysis derives invariant ordering at considerable cost, because it completely fails to take into account the agglutinative structure of clitic clusters[59].

As is well known, most clusters exhibit transparent concatenated sequences of clitics. Let us briefly survey some of the properties found in EP which suggest that clusters should be analysed as concatenated sequences of affixes.To start with, an 'opaqueness' approach fails to capture that, in strictly agglutinating clusters, clitics combine productively with other clitics, as in the case of EP *se* which can combine with six different dative clitic forms, as in (23). In Monachesi's analysis, however, each cluster is regarded as an individual unit, failing to capture the combinatorial possibilities allowed by the system and treating the similarity between clitic clusters with similar elements, as in (23), as a purely accidental phenomenon.

(23)

se + { me, te, lhe, nos, vos, lhes }

Then there are cluster-internal morphophonological effects which are exactly identical to those found at the verb-enclitic boundary. As alluded to before (cf. 9.2.2 and Chapter 4), 3rd person accusative clitics *–o*, *-a*, *-os* and *-as* become *–lo*, *-la*, *-los*, *-las* whenever they are preceded by a consonant-final verb (e.g., *levas-o* →*leva-lo*). In Chapter 7, I have shown that the verb-enclitic

[59] This problem is also found in syntactic analyses of cliticisation which adopt the idea that clusters are lexically stored portmanteau units (e.g., Schwarze 2001, Vigário 1999b).

context can be derived through morphophonological rules which select the appropriate accusative allomorph and trigger deletion on the verb. I have also mentioned that the same *l*-form of the accusative appears inside the clitic clusters whenever the accusative form is preceded by a consonant-final clitic (e.g., *nos-o* → *no-lo*). To me this suggests that there is only one morphophonological rule which applies whenever a 3rd person accusative allomorph is preceded by a consonant-final verb or clitic. However, this generalisation can only be captured if multiple occurrences of clitics are analysed as concatenated affixal sequences.

Also the set of combinatorial constraints which is typical of clitic clusters does not seem to prevent us from analysing clusters as sequences of affixes. In EP, the fact that the reflexive-dative cluster can only combine third person reflexives (cf. 23) or that dative pronouns only combine with first and second person accusative forms can evidently be captured by either formulating co-occurrence restrictions or, as suggested in this thesis, by formulating rule blocks so as to prevent these clitics from co-occurring (cf. Chapter 6 and below). These restrictions of co-occurrence therefore do not suggest that two-position clusters should be derived as phonologically opaque units. Finally, there are only two available orders, namely reflexive before dative, and dative before accusative, but again these ordering restrictions can be easily captured through rule block order, as shown in Chapter 6 for clitic suffixes.

Viewing all clusters simply as portmanteau units will not only fail to capture the crucial distinction between transparent sequences of pronouns and genuinely opaque ones, but also fail to provide an account for the cluster-internal regularities. Indeed, the only cases of portmanteau clusters in EP, for example, are given in (24) where dative clitics combine with 3rd person accusative clitics (cf. Chapter 4 for glosses).

(24)

a.	*mo* (me+o) / *mos* (me+os)	a'. *ma* (me+a) / *mas* (me+as)
b.	*to* (te+o) / *tos* (te+os)	b'. *ta* (te+a) / *tas* (te+as)
c.	*lho* (lhe+o) / *lhos* (lhe+os)	c'. *lha* (lhe+a) / *lhas* (lhe+as)

These clusters may constitute the exception to what I have said so far about the transparency of clitic sequences. But even in this case it may be argued that a simple allomorphic rule can derive the clusters in (24) by deleting the final vowel of the dative clitic (cf. Chapter 7).

Even though this analysis treats clusters as opaque morphological units, it provides an interesting solution for clitic ordering because it assumes that proclitic and enclitic clusters are not generated as separate sequences. We shall examine a similar proposal in Anderson (in press) below within Optimality Theory (Prince&McCarthy 1993). Note however that within the theory of Paradigm-Function Morphology, as developed by Stump (1993, 2001), preverbal sequences of affixes can only be generated as sequences of prefixes, i.e, as sequences of affixes that attach on a one-by-one basis to a preceding anchor point.

ANDERSON (1996, IN PRESS)

One further inflectional analysis concerned with the derivation of invariant clitic order is Anderson (1996, in press) who shares with Monachesi the view that cluster formation and cluster placement must be treated as somewhat independent stages of cliticisation. For Anderson, however, clusters constitute concatenated sequences of affixes (cf. also Miller and Sag 1997, Spencer 2000). So, unlike Monachesi (1999), who regards two-position clusters as opaque portmanteau units, it is possible for Anderson to generalise over cluster-internal regularities (cf. Legendre 2000). For Anderson, then, the question is how to linearise each clitic individually within the cluster and at the same time ensure that the order remains invariant in proclitic and enclitic position.

To accomplish this goal, Anderson assumes that regardless of where the cluster appears with respect to the verb, there is only one set of left-alignment constraints for the derivation of cluster formation. Within Optimality Theory (Prince&McCarthy 1993), each clitic is treated as an individual affix and is subject to its own set of alignment constraints. The cluster is derived under the assumption that the clitic with the highest ranked constraint will be leftmost, while all the other

clitics will be positioned to the right of the leftmost clitic. The verb is also assumed to be subject to left-alignment constraints. For proclisis, which is regarded as the default placement, the verb ranked lower and follows the cluster; for enclisis, the verb is ranked higher and appears as the leftmost element, to the left of the cluster. In either case, the same family of left-alignment constraints derives the order of clitics inside the cluster regardless of where the verb appears.

What seems crucial is that both Monachesi (1999) and Anderson (in press) assume, in different ways, that there is in effect only one set of constraints for the derivation of cluster formation. This insight has two very important results: a) it accommodates the invariant order of clusters and b) it captures the fact that the cluster appearing in proclitic position is exactly identical to the cluster appearing in enclitic position.

One problem with Anderson's analysis is that clitic linearisation takes place, in effect, in the syntax. As mentioned before, clitics in his analysis are regarded as phrasal affixes, i.e., affixes which attach to a syntactic domain. If clitics are linearised one by one with respect to a syntactic domain (not with respect to the verb), then clitic clusters are effectively generated as a sequence of phrasal affixes. Monachesi (1999), in this respect, generates the cluster in the morphology before attaching it to the verb. Even though no morphological structure is assigned to the cluster, it constitutes nonetheless a purely morphological unit. On the contrary, it is not clear how allomorphic effects and fusional morphology would be derived under Anderson's phrasal concatenation of clitics (cf. also Legendre 2000 for a similar account of cluster formation).

9.2.3 Summary

Particular emphasis has been given to the fact that the order of clitics inside the cluster, in EP and in Romance in general, remains invariant regardless of whether the cluster appears in proclitic or enclitic position. A tentative analysis of clitic order was proposed which showed that the standard derivation of prefixes and suffixes, as assumed within Paradigm Function Morphology (Stump

2001), derives preverbal clusters as mirror images of enclitic clusters and misses important generalisations about allomorphic alternations and co-occurrence restrictions (9.1.2). This lead us to examine alternative views of cluster formation which do not regard the one-by-one attachment of affixes as a mandatory principle. Instead, these studies share the insight that sequences of clusters should be regarded as morphological wholes with the ability to be placed as enclitic or proclitic sequences (Gerlach 1997, Monachesi 1999, Spencer 2000, Anderson 1996). Enclitic clusters are not viewed as sequences of suffixes nor are proclitic clusters generated as sequences of prefixes. Instead, clusters are neither inherently suffixal nor inherently prefixal, but simply 'self-contained' morphological wholes (9.1.3).

In the remainder of this chapter, I shall attempt to capture this insight within the theory of Paradigm-Function Morphology. I shall assume with Anderson (in press) that clusters constitute concatenated sequences of affixes (cf. also Miller&Sag 1997, Spencer 2000, Legendre 2000) and with Monachesi (1999) that clitic clusters are generated as morphologically cohering units (cf. also Miller&Sag 1997, Spencer 2000).

9.3 Realisation and placement of clitic clusters

This section shall attempt to accommodate rigid clitic ordering within Paradigm Function Morphology. A modified version of the theory shall be proposed, based on Spencer (2000), Luís (2001a, 2003b, to appear), Luís&Spencer (in press) and Spencer (ms). It explores the idea that clitic clusters are derived as whole units prior to their placement in proclitc or enclitic positon. This view is shared with previous inflectional and non-inflectional accounts, such as Anderson (1995), Monachesi (1999), Simpson&Withgott (1986), Grimshaw (1995), Gerlach (1997), Crysmann (2002), Spencer (2000). The main contribution of this section will be to show that this insight can be captured within the realisational model of Paradigm-Function Morphology. The amin idea is that

Stump's treatment of dual affixes in Swahili can be extended to clitic sequences in order to account for cluster formation.

9.3.1 Introduction

To begin with, there is no obvious way in which to accommodate within Paradigm-Function Morphology (Stump 2001) the insight that cluster formation and cluster placement should be regarded as separate stages of cliticisation. Possible routes shall now be considered, including Spencer (2000) for Bulgarian.

Let us imagine, for the sake of the argument, that a clitic sequence such as *no-lo,* as derived in Chapter 6, could be somehow 'cut off' from the verbal base and placed in proclitic position. Ignoring the domain of placement (cf. 9.2.2), this proposal would correctly capture that the order inside the cluster remains invariant. In addition, it would also use only one set of constraints for the derivation of the cluster and assume that its placement as a proclitic sequences does not affect the derivation of the clitic cluster.

This approach of course is conceptually untenable, because of the way in which realisation rules standardly apply in Paradigm-Function Morphology. As alluded to before, the format of RRs (cf. 6.3) derives each affix with respect to an anchor point (also known as input) which is generally an underived morphological base to which the affix attaches. Each affix therefore forms a *cohering unit* with the base and yields a form which in itself serves as input to the following affix (see cyclic application in 9.1.1). Once these suffixes are realised, there is no obvious way in which one might regard them as a separable sequence. Instead, an inflected verb is best seen as a layered structure. Layeredness here does not mean that the structure is hierarchical - it simply means that RRs apply sequentially to the output of the previous rule (Anderson 1992, Stump 2001). So, once a cliticised verb such as *entregava-se-lhe* is fully inflected, we cannot detach the string of affixal clitics (e.g.,

se-lhe) from the verbal stem. I shall therefore assume that enclitic clusters cannot be generated as genuine sequences of suffixes.

Let us now consider an alternative approach. Spencer (2000) provides the first treatment of cliticisation within Paradigm-Function Morphology and addresses the problem of cluster formation in Bulgarian. Spencer assumes that clitic clusters are derived inside the morphology as sequences of morphological units, like Monachesi (1999) and Miller&Sag (1997), and proposes that once clitic clusters have been derived, they can be placed as phrasal clusters. Let us now focus on the details of cluster formation and postpone the discussion about placement until 9.2.3.

Clitic clusters are generated as pseudo-suffixes, under the assumption that the leftmost clitic inside the cluster, in this case, cl_1 in (26), attaches to a null entity ∅. The output of attaching cl_1 to ∅ yields A which in turn serves as input to cl_2 in (26b). Applying cl_2 to A yields B, and so forth:

(26) a. cl_1: [∅ ---] A
b. cl_2: [[ ]A ---] B
c. cl_3: [[ ]B----]C

Within this approach, the sequence of clitics is generated within the morphology as a 'disembodied' cluster, i.e., a sequence of affixes without an ordinary morphological anchor point, and placed in the syntax as a phrasal affix (cf. also Halpern 1995 for Serbo-Croatian). Unlike Anderson (in press) and Legendre (2000) (cf. 9.1.3), clusters are not concatenated in the syntax but inside morphology. Once clusters have been formed, placement rules can apply and determine where clusters should appear.

The problem with Spencer's analysis is that it has been conceived for a language in which clusters attach systematically to the right of a given host, regardless of whether they are enclitic or proclitic with respect to the verb (Halpern 1995). In this respect, the null entity in (26a) works as a 'place holder' for the host. In Romance, however, clusters attach not only leftwards (for enclisis) but

also rightwards (for proclisis). The idea that there is a null entity at the left edge of the clusters serving as the 'anchor point' might still work when clusters are positioned enclitically; but it would be counterintuitive to make the same assumption when the clitic cluster attaches as a proclitic sequence. The proposal made by Spencer (2000) can therefore not be adopted for the case of the clitic systems with dual placement of clusters.

Instead, for Romance, clitic sequences need to be analysed as inflectional units without an anchor point and assigned the ability to appear before or after a clitic host, as proposed in Monachesi (1999). This idea is illustrated informally below:

(27) cluster: *se-lhe*

(28) cluster placement: enclisis <X , <*se-lhe*> >
proclisis << *se-lhe*>, X>

The problem with this view is that it cannot be captured within the present model of the theory of Paradigm-Function Morphology (Stump 2001).

9.3.2 Proposal

To be able to accommodate the distinction between (27) and (28) within Paradigm-Function Morphology, it will be necessary to treat cluster formation and cluster placement separately. Let us therefore return to Stump's 'separationist' treatment for the dual placement of affixes (section 9.1) and consider how it might be extended to clitic clusters.

a) separating exponence from placement

As summarised in section 9.1, the 'separationist' approach to inflection (Stump 1993) offers a treatment for the dual placement of affixes. The analysis is based on the assumption that the form and placement of affixes can be captured separately. Realisation rules are formulated which do not

specificy the direction of attachment of an affix; instead, whether affixes are attached as prefixes or suffixes is determined by separate placement rules which specify i) the conditions under which affixes are positioned and ii) their position with respect to the morphological base. Let us therefore assume, first of all, that Paradigm-Function Morphology allows us to formulate placement constraints that are independent of the rules realising affixal clitics.

Let us now explore the 'separation-approach' to clitic sequences. Note that clusters are also 'dual' entities, in the sense that one and the same clitic sequence can be positioned before or after the clitic host. Recall from section 9.2, that Stump's proposal cannot be adopted in its entirety because affixes are attached sequentially, that is, the output of attaching an affix from one block becomes the 'base' for the affix in the next block (cf. section 9.1.2).

Nonetheless, Stump's insight about the separation between the form and the position of affixes may help us accommodate the dual placement of clusters. What we need to assume is that rather than defining placement for each affix, we only define it once for the whole clitic cluster (cf. 27-28). This view entails that the cluster must be generated prior to its placement.

In what follows we will consider the analysis of cluster realisation.

b) clusters as whole inflectional units

If we are to generate clitic sequences separately from their host we need to use a set of notational conventions for realisation rules that are different from those of standard Paradigm-Function Morphology. In particular, we need to formulate rules that do not operate recursively[60], that is, they are designed to feed each other. Let us therefore explore the idea that realisation rules (RRs) are

[60] Standard rules in this context refer to both ordinary realisation rules (deriving ordinary suffixes, cf. ch.6) and 'dual/ambifixal' inflectional rules (proposee in Stump 1993). Both of them apply sequentially.

functions and make use of function composition. Consider the hypothetical Paradigm Function in (29) in which standard RRs apply cyclically, where the first rule feeds into the next, and so on.

(29) $PF(X, \sigma) =_{def} RR_z(RR_y(RR_x(X, \sigma))$

As Spencer (2000) points out, it is possible to think of such a string as a single, composite function, by virtue of function composition. This means that the layered representation of RRs in (28) is identical to the composed representation in (30).

(30) $PF(X, \sigma) =_{def} (RR_x \circ RR_y \circ RR_z) (X, \sigma)$

Through function composition, the hypothetical string of affixes generated by the realisation rules RRx, RRy and RRz can be derived as a composed unit ready to be placed (as a whole) to the left/right of the host. So, rather than positioning affixes on a one-by-one basis, the placement function will be defined only once for the whole affix string.

To clarify the assumptions underlying this view, let us suppose that we have two functions, f(x) and g(x) and suppose they have the same domain and range (for instance, they might be arithmetic functions over numbers). Then the value of f is itself a number and in the domain of g. Thus, we can define g over the output of f: g(f(x)). For instance, suppose that f is 'multiply by 10' and g is 'divide by 5'. Then f(7) = 70 and g(70) = 14. But we can conflate the two operations into a single operation of 'multiply by 10 then divide by 5': g(f(x)). This conflated function is the composition of the two functions, denoted 'f◦g'. In other words, we can say (f◦g)(7) = 14.

Exactly the same can be done with realisation rules: since realisation rules are functions it is a trivial matter to define the composition of a set of realisation rules (Stump 1993, 2001, Chapter 5; Spencer 2000, Luís&Spencer in press). Following this set of assumptions about function composition, the clitic clusters shall be derived precisely as a composed sequence of affixal clitics whose placement is determined by a set of placement rules (cf. 9.2.3.3).

Although we need to use different notational conventions from those of standard PFM, this proposal appears to be the logical outcome of the arguments provided in this section, since there is simply no other way of deriving the data without relaxing the idea that affixes must attach sequentially (Luís 2001a, 2003a). In the following section, I shall illustrate how this proposal can account for cliticisation in EP, including other Romance languages.

9.3.3 Cluster formation and placement

This section provides an explicit analysis of cluster formation and cluster placement within a modified version of the theory of Paradigm-Function Morphology (PFM), making use of two crucial ideas: namely that a) exponence and placement should be derived separately, and that b) affixal sequences should be derived through function composition.

First, I shall revise the realisation rules proposed in Chapter 6 for affixal clitics and illustrate how function composition applies to EP clitic pronouns (9.2.3.1). Placement constraints shall be formulated and special emphasis will be given to the different domains within which clusters can attach (9.2.3.2). I shall then consider how cliticisation interacts with ordinary affixation, based on the discussion offered in Chapter 6 (9.2.3.3).

9.3.3.1 Generating affixal strings

To capture the dual placement of affixal clitics, a format for realisation rules must be adopted which is different form the one adopted for ordinary verbal suffixes, such as tense and agreement. For the latter, I shall maintain the standard assumption (Stump 2001) that they are defined through ordinary realisation rules (RRs) (cf. Chapter 6). These are rules which apply sequentially to the stem of a given lexeme. This means that affixes are realised and attached on a one-by-one basis and that RRs conflate exponence with placement. For affixal clitics, on the contrary, a different set of RRs must be used in order to separate the exponence of clitics from their placement. I shall follow Stump's

analysis of Swahili dual affixes and formulate RRs which define only the form of the affix. The format of these rules is given below.

(31) $RR_{n, \tau, C}$ (<X, σ>) = <aff, σ>

As mentioned before, each RR carries three indices: n is the rule block index, τ is the feature set which the rules realise, and C is the set of word class features over which the rule is defined. The notation < *affix,σ* > is interpreted as referring to an affix which realises a given set of morphosyntactic features and whose linear order with respect to the host is unspecified. Following Stump (1993), placement is stated separately by a placement function as shown below (cf. 9.2.3.2). This means that the base X over which the RR is defined is therefore not part of the output of the rule, unlike in the standard format of RRs (cf. Chapter 2).

Given this new format of RRs, I suggest that the rules previously formulated in Chapter 6, based on the ordinary format, should be rewritten as follows (cf. Appendix B)[61]:

(32) Clitic rule blocks

a. Block A

i. $RR_{A1, \{Person:3, Refl:+\}, V}$ (σ) $=_{def}$ [se, σ]

b. Block B

i. $RR_{B1, \{Person:1, Number: Sg\}, V}$ (σ) $=_{def}$ [me, σ]

ii. $RR_{B2, \{Person:2, Number: Sg\}, V}$ (σ) $=_{def}$ [te, σ]

iii. $RR_{B3, \{Person:1, Number: Pl\}, V}$ (σ) $=_{def}$ [nos, σ]

iv. $RR_{B4, \{Person:2, Number: Pl\}, V}$ (σ) $=_{def}$ [vos, σ]

v. $RR_{B5, \{Case: Dat, Person:3, Number: Sg\}, V}$ (σ) $=_{def}$ [lhe, σ]

vi. $RR_{B6, \{Case: Dat, Person:3, Number: Pl\}, V}$ (σ) $=_{def}$ [lhes, σ]

[61] For simplicity, I represent the affixes in their default form. However, I assume that realisation rules also encode morphophonemic information about the affixes they realise (Stump 2001: 45), as proposed in Chapter 7 (cf. Appendix B).

c. Block C

i. $RR_{C1,\ \{Case:\ Acc,\ Person:3,\ Number:\ Sg,\ Gender:\ Masc\},\ V}\ (\sigma) \quad =_{def}\ [o, \sigma]$

ii. $RR_{C2,\ \{Case:\ Acc,\ Person:3,\ Number:\ Sg,\ Gender:\ Fem\},\ V}\ (\sigma) \quad =_{def}\ [a, \sigma]$

iii. $RR_{C3,\ \{Case:\ Acc,\ Person:3,\ Number:\ Pl,\ Gender:\ Masc\},\ V}\ (\sigma) \quad =_{def}\ [os, \sigma]$

iv. $RR_{C4,\ \{Case:\ Acc,\ Person:3,\ Number:\ Pl,\ Gender:\ Fem\},\ V}\ (\sigma) \quad =_{def}\ [as, \sigma]$

Compared to the analysis in Chapter 6, the same number of rule blocks is maintained and the featural description of each realisation rules is also in conformity with the earlier proposal. The notational changes that were introduced simply show that the rules in (32) do not have the abilty to determine whether the attachment of affixes is prefixal or suffixal. This means that proclitic and enclitic clitics are generated through the same set of RRs.

Note that even though we can now predict that enclitics and proclitics are exactly identical in form, number and meaning, we must still prevent each affix from attaching individually to the host. In other words, we need to ensure that placement is defined only once over the whole clitic sequence.

It shall therefore be assumed that the RRs in (32) are combined through function composition in order to derive clitic clusters as a 'composed' string of affixes unspecified for placement. In (33) the string of affixes <*aff-aff*> constitutes a composed unit ready to be attached to a host.

(33) $(RR_A \circ RR_B \circ RR_C)\ (\sigma_2) =_{def}\ [\ <aff\text{-}aff>,\ \sigma_2]$

To illustrate, suppose we wish to generate the cluster *se-lhe* 'himself/herself to him/her'. Based on the individual rules given in (32), the composed function in (33) will be evaluated as in (34)[62].

(34) $(RR_{C\{\}} \circ RR_{B5} \circ RR_{A1})\ (\sigma_2) =_{def}\ [\ <se\text{-}lhe>,\ \sigma_2]$

[62] As already mentioned, only at most two affixal slots are filled in any given cluster, however all three clitic rule blocks must be defined in the PF (cf. Chapters 2 and 6).

Finally, even though there is no successive rule application, in the traditional sense, it is crucial to ensure that the correct clitic order is derived, namely that reflexives precede datives and that datives precede accusatives. This is done by rule blocks and rule block indices which define the linear order of clitics inside the string. So, the fact that the notion of rule block and rule block index on realisation rules is maintained, as shown in (33), enables us to correctly capture clitic order presented in Chapter 4. In this sense, I assume that RR_B in (35) corresponds to rule block II and that RR_A to rule block I; this means that RRI applies before RRII, correctly deriving the sequence *se-lhe*.

9.3.3.2 Cluster placement

Having offered an explicit characterisation of cluster formation, we must now define cluster placement. As mentioned above, through function composition the clitic cluster is derived as a composed unit ready to be positioned to the left or right of the clitic host (ignoring for now the exact identity of the host). I shall assume that independent placement rules, as shown in (35), determine whether the cluster is proclitic or enclitic.

(35) Placement rules (provisional)

{<aff,aff>, X} →

a. Proclisis: < <aff,aff>, X> under certain conditions

b. Enclisis: < X , <aff,aff> > , under certain conditions

So, rather than positioning affixes on a one-by-one basis, the placement function will be defined for the whole affix string. Two placement rules are provided: (35a) positioning the cluster before the host and (35b) placing it after the host. Using this approach, we solve the rather complex problem discussed earlier in this chapter about the derivation of enclitic and proclitic clusters: we capture that clitic order remains invariant and we accommodate the ability for a given string to appear either as an enclitic sequence or as proclitic sequence.

We need however to be slightly more specific with respect to these rules. Note that as formulated above, they tell us very little about the identity of the host and the conditions under which each placement takes place. Both these 'parameters' however constitute essential information for characterising cluster placement.

a) factors triggering placement

In most Romance languages, excluding EP, cluster placement is generally conditioned by the finiteness properties of the verb. This generalization holds for languages such as Spanish, standard French or Italian in which proclisis is triggered by finite verb forms, while enclisis is selected by non-finite or imperative verb forms (Nishida 1987, Brines 2000, Miller&Sag1997, Monachesi 1999, among others). Assuming that cluster placement in these languages is proclitic by default, we can formulate the placement constraint as in (36).

(36) Placement rules for Spanish, standard French, Italian

a. Enclisis: < X , <aff,aff> >, if verb form non-finite or imperative.

b. Proclisis: < <aff,aff>, X> , elsewhere.

The case of EP is different for two reasons. First, the alternation between enclisis and proclisis is not affected by the finiteness properties of the verb, but by syntactic properties of the clause (cf. Chapters 8 and 10). To capture the effect of the syntax, it will be suggested that proclitic placement is found solely with the verb forms associated with a formal morphological feature [Restricted: yes] which is linked in the syntax to the contexts triggering proclisis (Luís&Sadler 2003). The [Restricted] feature is just one of a number of technical devices one might propose to deal with the interaction between the syntax and the morphology. What we are effectively saying is that, for the verb forms bearing this feature, the placement function will be governed by syntactic principles. A more detailed discussion about the arguments underlying the use of this formal feature is provided in Chapter 10, section 10.3.

The second difference between EP and the other Romance languages is that the default position of the cluster is enclitic (Duarte et al 1995). Supporting this idea is data from language acquisition and from adult speech, which show that children produce enclisis at an earlier stage and that adults tend to use enclisis even in proclitic contexts.

To capture both a) the conditions triggering the alternation between enclisis and proclisis, and b) the default status of enclisis, cluster placement will be specified as in (37).

(37) Placement rules for EP (provisional)

{<aff,aff>, X}→

a. Proclisis: < <aff,aff>, X >, if verb form [Restricted: yes]

b. Enclisis: < X, <aff,aff> >, elsewhere

b) Identity of the host

In previous sections I addressed the need to define the identity of the clitic host. Recapitulating what I said earlier, we need to state that the identity of the host varies within Romance and even within one given language (e.g., EP). For Spanish, for example, the host is uniformly morphological. For EP two domains of clitic placement are necessary, a phrasal domain for clitic prefixes and a morphological domain for clitic suffixes, as shown in (38). Thus, enclitic clusters attach to a 'stem' (cf, Chapter 6) and proclitic clusters to a phrasal Vº node, immediately dominated by VP (cf. Chapter 8):

(38) Placement rules for EP

{<aff,aff>, X}→

a. Proclisis: < <aff,aff>, Vº-VP>, if Vf [Restricted: yes]

b. Enclisis: < stem, <aff,aff>, elsewhere

This study therefore is able to capture cases of uniform cluster placement, as in Italian or Spanish, where clitic clusters attach to a morphological host (Monachesi 1999, Brines 2000), but it can also

accommodate cases of asymmetric placement as in EP and capture the co-existence within a given language of both word-level and phrasal attachment. On the contrary, Crysmann's (2002) account of EP is unable to captured the fact that strong phonological and affixal similaritites between EP clitics and clitic pronouns in other Romance languages. By introducing a special 'hybrid' category to treat EP clitic pronouns, his analysis suggests that clitic pronouns in EP are categorially different from clitic pronouns in languages such as Spanish and Italian.

9.3.3.3 Revised Paradigm Function

Given the above proposal about cluster derivation and placement, we now need to slightly revise the Paradigm Function proposed in Chapter 6 for cliticised verbs. As I will show next, the alterations only affect the clitic layer, given that I am assuming that ordinary tense and agreement suffixes are realised in a standard fashion, i.e., through realisation rules which conflate exponence and placement[63].

In Chapter 6, it was argued that the paradigm function defining cliticised verbs operates over two morphosyntactic feature sets: one set σ_1 stands for the verbal features and σ_2 for clitic features. This division should capture that cliticised verbs comprise both ordinary verbal features (e.g, tense and subject agreement) and clitic features (e.g., case, object number and object person). In an ordinary PF the complete set of morphosyntactic features is realised by one group of rule blocks (cf. Chapter 2). The division of morphosyntactic features therefore means that the realisation rules generating the cliticised verb form are split up into two sets of rule blocks: RRs deriving ordinary verbal suffixes and RRs deriving the clitic string.

---

[63] Even though I formulate the 'separation' hypothesis with respect to affixal clitics, there is no technical impediment for applying the idea to all affixation. Such a 'generalised' use of the 'separation' is adopted by Spencer (ms) and Luís&Spencer (in press).

The revised PF however is different from the one proposed in Chapter 6 in that two tense and agreement suffixes are realised through standard RRs (39a), while the clitic cluster in (39b), on the contrary, is derived in two stages: as a composed unit through function composition (39bi) and attached to a clitic host through a placement function (39bii).

(39) Revised Paradigm Function for cliticised verb forms

Where X = stem of *L* and $\sigma = \sigma_1$ (= verbal features) and σ_2 (= clitic features),

PF (<X, σ>)

$=_{def}$ { a. base { i. stem: X- ; ii. verbal suffixes: RR_{II} (RR_I (σ_1) } ; b. cluster { i. exponence: ($RR_A \circ RR_B \circ RR_C$) ($\sigma_2$) ; ii. placement { < <aff,aff >, V^o >, if Vf [Restricted: yes] ; stem, <aff,aff > >, elsewhere } }

As argued in Chapter 6, the general format of the paradigm function thus captures that cliticised verbs comprise two layers of inflection.

9.3.4 Enclisis and mesoclisis

This section provides sample derivations of enclisis, mesoclisis and proclisis in EP. The analysis is based on the proposals made in Chapter 6 about 'two-layered' Paradigm-Functions and the proposal made in this chapter about cluster formation. For the derivation of tense and agreement suffixes, I adopt the realisation rules introduced in Chapter 6 (cf. also Appendix A); clitic clusters shall be derived through the realisation rules in (32), also summarised in Appendix B.

9.3.4.1 Enclisis

For the verb form *apresentavam-se-lhe* ‘introduced himself/herself to him/her’, the PF in (40) might be evaluated as follows:

(40) PF-analysis of *apresentavam-se-lhe* 'they introduced themselves to him/her'

Where σ_1 = {Tns:Impf, Per:3, Num:PL}

and σ_2 ={[Case:Ref,Per:3, Num:Pl]; [Case:Dat,Per:3,Num:Sg]},

PF (<apresenta-, σ>)

a. $=_{def}$ { base { stem$_1$ of L: apresenta-; verbal suffixes: RR_{IId} (RR_{Ia} (σ_1) }; cluster { exponence: ($RR_{A1} \circ RR_{B5} \circ RR_{C\{\}}$) ($\sigma_2$); placement: <stem, <aff,aff> >, elsewhere }

b. $=_{def}$ { base: apresenta-va-m; cluster { exponence: ($RR_{A1} \circ RR_{B5} \circ RR_{C\{\}}$) ($\sigma_2$); placement: <stem, <aff,aff> >, elsewhere }

c. $=_{def}$ { base: apresentavam; cluster { se-lhe; <stem, <aff,aff> >, elsewhere }

d. $=_{def}$ <apresentavam-se-lhe, σ>

In (42), the first two lines specify the feature set σ_1 and σ_2 associated with the cliticised verb *apresentavam-se-lhe*. The evaluation proceeds as follows: in (40a), the PF is defined over the verbal stem *apresenta-* of the lexeme APRESENTAR, and the tense and agreement affixes are assumed to apply to this 'BASIC' stem (i.e., stem$_1$) through the piecewise application of RR_{Ia} and RR_{IId}. At this stage of the derivation, tense and agreement suffixes are combined with the stem to yield the clitic base *apresentavam* in (40b). So far, the analysis follows the proposal from Chapter 6.

Let us now turn to the clitic clusters. Whereas the cluster was derived as a genuine suffixal sequence in Chapter 6, it will now be derived as a composed unit. Rule block A realises the clitic *se* and block B the clitic *lhe* (40c). Rule block C supplies no exponent and must be evaluated through the identity function[64]. Rule block order ensures that the the clitic sequence is linearised inside the cluster, and function composition ensures that the linearised sequence is derived as a composed unit. The placement constraint "<stem,<aff,aff>>" attaches the clitic cluster to the right edge of a verbal stem form to yield the cliticised verb *apresentavam-se-lhe* in (40d). As argued before, placement takes place only once, after the cluster has been generated.

9.3.4.2 Mesoclisis

For mesoclisis, the placement of clitic clusters with respect to verbal suffixes is reversed. As proposed in Chapter 6, we formulate a constraint which changes the order in which clitics and ordinary suffixes appear. This special rule is sensitive to the {Future/Conditional} features and will allow the clitic features to be applied directly to the future and conditional stem form (i.e., $stem_2$, cf. Chapter 6). The revised format of the rule is given below:

(41) Paradigm Function for cliticised verb forms with verb-internal clusters

Where σ_1 = {Tns: Fut/Cond} and σ_2{Case: α; Person: β, Number: γ},

PF (<X, σ>)

$=_{def}$ { base { i. $stem_2$ of L X-; ii. cluster { exponence $(RR_A \circ RR_B \circ RR_C)(\sigma_2)$; placement <X, <aff,aff > > } }; verbal suffixes $RR_{II}(RR_I(\sigma_1))$ }

[64] As mentioned before (cf. Chapter 6), this is a function which when applied to an argument delivers that argument as its value: f(X) = X. In other words, the input to $RR_{C\{\}}$ is also its output.

By reversing the order in which each one of the so-called affix-layers is realised, we capture the fact that clitic clusters occur verb-internally under specific conditions. A sample derivation of mesoclisis is given in (42) for the verb form *levar-lhe-emos* 'we will bring to him/her'.

(42) PF-analysis of *levar-lhe-emos*

Where σ_1={Tns: Fut, Per: 1, Num: PL} and σ_2 ={[Case: Dat, Per:3, Num:Sg],

PF (<levar, σ>)

a. $=_{def}$ base { stem$_2$ of L: levar- ; cluster { exponence: ($RR_{C\{\}}$ RR_{B5} $RR_{A\{\}}$) (σ_2) ; placement: <X, <aff,aff > > } ; verbal suffixes: RR_{IIb} (RR_{Id} (σ_1)

b. $=_{def}$ base { stem$_2$ of L: levar- ; cluster { exponence: lhe ; placement: <X, <aff,aff > > } ; verbal suffixes: RR_{IIb} (RR_{Id} (σ_1)

c. $=_{def}$ base: levar-lhe ; verbal suffixes: RR_{IIb} (RR_{Id} (σ_1)

d. $=_{def}$ <levar-lhe-emos, σ>

If we follow the path *base-cluster-exponence&placemen,* we observe that the cluster in (42a) is defined both in terms of form and placement. Assuming that affixal sequences are derived through function composition, the cluster will be realised as a composed unit; ti will also be associated with a postverbal alignment constraint which positions the cluster immediately after the infinitival stem. Like enclisis, then, mesoclisis constitutes clitic suffixation. The difference between one and the other resides in the order in which the ordinary layer and the clitic layer occur.

In the derivation of the cluster, both rule blocks A and C are realised vacuously because no explicit rule from these blocks can apply. By the Identity Function Default, then, no exponent is supplied. Only rule block B supplies an exponent, as given in (42b). This clitic combines with the infinitival stem *levar-* yielding a cliticised infinitival stem (42c). The tense and agreement suffixes are realised sequentially and attached one-by-one to the cliticised stem *levar-lhe* yielding the cliticised form *levar-lhe-e-mos* in (42d).

Given this analysis, mesoclisis is effectively derived as a type of metathesis and the verb forms as a morphologically synthetic word. As mentioned in Chapter 6, the crucial point in the derivation of mesoclisis is to capture that the verb-internal clitic sequence attaches to an infinitival stem (before the tense and agreement marker). From the combination of this stem with the cluster a morphologically more complex base is derived which serves as input to the deriving ordinary verbal suffixes.

9.3.4.3 Morphophonology revisited

Chapter 7 sketched an inflectional analysis of clitic-induced allomorphy based on the combination of morphophonological rules (m-p rules) with metagenaralisations. While m-p rules constrain the evaluation of one or more realisation rules (cf. 7.3.2.1), metagenaralisations associate m-p rules with the appropriate realisation rules (cf.7.3.2.2). In this short section, I will show that the separation between exponence and placement, proposed in this chapter, for clitic clusters and cliticisation does not hinder us from applying the set of m-p rules and metageneralisations introduced earlier on.

Let us first recall some of the m-p rules and the corresponding metageneralisations. Two of the m-p rules are repeated in (43) and the metageneralisations are given in (44).

(43) Fragment of m-p rules (cf. (28) in Chapter 7)

Where $RR_{n, \tau, C}(X, \sigma) =_{def}$ $<Y', \sigma>$:

a. If X = Wv[nasal] and Y = XLNZ, then Y' = XnZ
b. If X = W[consonant] and Y = XLNZ, then Y' = WlZ

(44) Fragment of metageneralisations (cf. (31) in Chapter 7)

a. m-p rule (45a) is associated with realisation Clitic-Block C, if σ_1 contains {Agr:3pl};
b. m-p rule (45b) … [is] associated with Clitic-Block C, everywhere.

(43a) and its associated metageneralisation (44a) apply to 3rd person accusative clitics whnever these are preceded by 3rd plural verb forms. Under this condition, (43a) derives the *n*-initial allomorphs of the 3rd person clitics (e.g., *levam-no* '(they) take him').

(43b), which is associated with the metageneralisation in (44b), applies to 3rd person accusative clitics whenever they are preceded by a consonant-final verb. The rule derives *l*-initial allomorphs and deletes the final consonant on the verb (e.g., *encontrámo-lo*, not **encontrámos-o* '(we) found him').

Two examples of allomorph selection are given below, using the partial derivations in (45-46). The verbal base in (45ci) serves as input to the rules realising and placing the clitic cluster in (45cii). The first line in (45cii) shows that the clitic cluster comprises only one clitic. By the RR rules proposed in chapter 7 (i.e., RRC1, see also Appendix B), the affixal clitic has been realised as a clitic exponent with underspecified phonology. The second line in (45cii) says that the clitic will attach after the verbal base. Given this adjacency requirement, the m-p rule (43a) applies prior to the placement of the clitic deriving the *n*-initial allomorph, as shown in (45d). A similar derivation is given in (46), except that here the m-p rule in (43b) applies, deriving an *l*-initial allomorph and triggering verb-final consonant deletion.

(45) partial derivation of *alimentavam-no* 'they fed him'

Where σ_1 = {Tns:Impf, Per:3, Num:PL;} and σ_2 = [Case:Acc,Per:3,Num:Sg; Gen: Masc]

PF (<alimenta-, σ>)

c. $=_{def}$ { i. base alimentavam; ii. cluster { LNo (by RR_{C1}); <stem, <aff,aff> >, elsewhere }

d. $=_{def}$ <alimentavam-no, σ>

(46) partial derivation of *alimentava-lo* 'you.sg fed him'

Where σ_1 = {Tns:Impf, Per:2, Num:Sg;} and σ_2 = [Case:Acc,Per:3,Num:Sg; Gen: Masc]

PF (<alimenta-, σ>)

c. $=_{def}$ { i. base alimentavas; ii. cluster { LNo (by RR_{C1}); <stem, <aff,aff> >, elsewhere }

d. $=_{def}$ <alimentava-lo, σ>

Both of the m-p rules that apply in (45) and (46) are associated with 3rd person accusative clitics. Recall that these clitics are realised with the abstract consonantal morphophoneme 'LN' (cf. Appendix B) capturing the fact that 3rd person accusative clitics have three allomorphs: l-type, n-type and the default. RR_{C1} realises the underlying LNo affix, while the m-p rule in (45a) evaluates it as an *n*-allomorph. The m-p rule in (45b) evaluates the affix as an *l*-allomorph.

This section illustrated the implementation of m-p rules under the assumptions that placement and exponence are separated: m-p rules are applied after placement has occurred. Most of the m-p rules are crucially sensitive to linear order and adjacency. Since RRs are still functions, the m-p rules can be associated (or tied) to each individual RR as in standard PFM.

9.3.5 Proclisis

Let us now examine how proclisis would be derived, given the assumption that preverbal clusters are attached through phrasal affixation. One of the claims argued for in the third part of this thesis is that EP has an asymmetric clitic system with two domains of attachment. A mixed clitic system then is simply a system in which affixal clitics can selelct between a morphological and a phrasal host. Uniform clitic systems would then be those in which affixal clitics attach only to morphological hosts (e.g., in Italian, Monachesi 1999; Macedonian, Spencer 2000) or to phrasal hosts (as in Bulgarian, as suggested by Spencer 2000, or in Kwakwala, as proposed in Klavans 1980, Anderson 1992).

9.3.5.1 Phrasal affixation

The phrasal placement of the cluster is determined by the placement constraint in (47). It says that the cluster is placed at the left edge of a Vº immediately dominated by VP, if the verb form to which the cluster attaches bears the feature [Restricted:yes]. Earlier I mentioned that proclitic placement is found solely with verb forms associated with a formal morphological feature [Restricted: yes] which is linked in the syntax to the contexts triggering proclisis (cf. 9.2.3.2; see also Chapter 10). By default, then, the feature Restricted is marked 'no' and placement is postverbal.

(47) Proclisis (cf. 38 in 9.2.3.2)
{<aff,aff>, X} → < <aff,aff>, Vº-VP>, if Vf [Restricted: yes]

A sample derivation of proclisis is provided in (48), based on the general Paradigm Function in (39).

(48) PF-analysis of *se-lhe apresentavam* 'they introduced themselves to him/her'

Where σ_1 = {Tns:Impf, Per:3, Num:PL}, [Restricted:yes] and
σ_2 ={[Case:Ref,Per:3, Num:Pl]; [Case:Dat,Per:3,Num:Sg]},

PF (<apresenta-, σ>)

a. $=_{def}$ { base { stem$_1$ of L: apresenta-; verbal suffixes: RR_{IId} (RR_{Ia} (σ_1)) }; cluster { exponence: ($RR_{A1} \circ RR_{B5} \circ RR_{C\{\}}$) ($\sigma_2$); placement: < <aff,aff>, V^o> } }

b. $=_{def}$ { base: apresentavam; cluster { exponence: ($RR_{A1} \circ RR_{B5} \circ RR_{C\{\}}$) ($\sigma_2$); placement: < <aff,aff>, V^o> } }

c. $=_{def}$ { base: apresentavam; cluster { e-lhe; < <aff,aff>, V^o> } }

d. $_{def}$ < se-lhe, apresentavam >

As in the previous derivation, the first two lines specify the feature set σ_1 and σ_2 that is associated with the cliticised verb. As in the case of verb-enclitic units, the PF in (48a) specifies that tense and agreement features are realised before clitic features; however, the placement function defined by the Pf for this verb form is < <aff,aff>, V^o>.

The evaluation then proceeds as follows: tense and agreement suffixes apply to stem$_1$ of APRESENTAR through the application of RR_{Ia} and RR_{IId} to yield the clitic base *apresentavam* in (48b). The clitic cluster then is linearised through rule block order and composed through function composition yielding *se-lhe* in (48c). Given the placement function < <aff,aff>, V^o>, we derive the cliticised verb form as in (48d). The cluster will be positioned in the syntax to the left edge of a V^o

node and therefore cannot be morphologically attached to the verb. Instead, it forms the verb a discontinuous predicate with the verb.

9.3.5.2 Discontinuous predicates

By enabling clusters to be positioned in two distinct components of grammar, we have defined two complementary verbal paradigms, one in which the verb combines with a suffixal sequence (cf. 9.2.4), the other in which the verb combines with a phrasal prefix (9.2.5.1). The former constitutes a morphologically cohering word form with the clitic string being part of verbal morphology; the latter on the contrary will be regarded as a discontinuous cell in the paradigm of the lexeme. By placing the cluster in the syntax, we are effectively saying that proclitics do not form a morphologically cohering unit with the verb, thus capturing the fact that they do not need to be adjacent to it, unlike enclitics.

One final word is in order about phrasal affixation. The proclitic-verb combination should not be confused with an analytical construction, as found with periphrastic constructions or separable-particle verbs, whose members constitute syntactic word level units (Ackerman&LeSourd 1997, Ackerman&Webelhuth 1998, Sadler&Spencer 2001, Brassil 2001, Kathol 1995, Lüdeling 2001, among others). Under lexicalist assumptions, the verbal base constitutes a syntactically independent word unit but the phrasal affix is regarded as 'floating' affix, one which is not represented under a syntactic node (cf. Chapter 8).

The claim that proclitic-verb combinations form discontinuous morphosyntactic expression appears to be the logical outcome of phrasal affixation. In Chapter 10, I shall briefly discuss the syntactic formalisation of phrasal affixation within the theory of Lexical-Functional Grammar (Bresnan 2001, Kaplan&Bresnan 1982). Further cases of phrasal affixation are summarised briefly in the following section.

9.3.6 Cross-linguistic evidence

One of the main claims of this thesis is that EP affixal clitics should be assigned the ability to be positioned either in the syntax or in the morphology (cf. Chapter 8). In this section, evidence suggests that this property is not exclusive to the EP clitic system. In languages such as Cappadoccian and Rumanian phonological evidence suggests that clitics also undergo both phrasal and morphological attachment (Popescu 2000, Monachesi 2000, Gerlach 2001a, Luís in press, 2003a). In Rumanian, it is the phonological behaviour of proclitics which suggests phrasal status; in Udi, it is the scopal behaviour of one set of subject agreement affixes which indicates that they cannot be morphologically part of the host. The data suggests that the Paradigm-Function analysis proposed for EP cliticisation may be adopted for other inflectional phenomena cross-linguistically.

a) Cappadocian (Janse 2000, Condoravdi&Kiparsky 2001)
Based on the Zwicky-Pullum criteria used in Chapter 4, it seems that clitic pronouns in Cappadocian (Asia Minor Greek dialect) display several affix-like properties. For example, they form rigidly ordered clusters (49a) and exhibit a high degree of selectivity both in preverbal and postverbal position (49a-b).

(49) a. léo se ta (Janse 2000)
I-say 2sg.dat acc.3pl.acc
'I will tell it to you'

b. itúta ta prámata vúla ta pírin
these the things all 3pl.acc he-took
'these things, he took them all'

In addition, enclitics share with verbal suffixes the fact that they induce non-productive phonological effects on the verb (Janse 2000). In particular, they can place an additional 'enclitic'

accent on a verb with antepenultimate stress (50b), and can shift stress to the penultimate syllable like suffixes (51b).

(50) a. éfayan (Janse 2000)
b. éfayán -do
they.ate -3pl.acc
'they ate it'

(51) a. sérepsan
b. serépsan -ta
they.gathered -3pl.acc
'they gathered them up'

As Janse (2000) observes, this type of stress shift takes place only between enclitics and verbs, suggesting that it takes place word-internally, rather than in phrasal phonology. Similar effects are found in Standard Modern Greek and used as evidence to support the claim that clitic pronouns are part of verbal morphology (Joseph 1988).

Further affix-like evidence is provided by the occurrence of morphophonological processes between the enclitic and the verb normally operating within words rather than across word-boundaries. A typical example is the voicing of the voiceless dental plosive /t/ after nasals (cf. Janse 1995) or assimilation processes as exemplified in (52).

(52) a. /rotún + čin/ rotún=dži (Janse 2000)
'they ask her'
b. /paγas+mes/ páaz=mes
'take us'

So far then, the evidence indicates that postverbal clitics are verbal suffixes.

Proclitics are formally exactly identical to enclitics and must be adjacent to the verb. This suggests that a) proclitics and enclitics constitute the same affixal unit and that b) proclitics are

placed at the left edge of the verb. There is however one crucial property of proclitics which indicates that they are not morphologically part of the verb: proclitics are phonologically enclitic to the preceding word.

As pointed out by Janse (2000), the phonological behaviour of modal particles shows that proclitics form a prosodic word with the preceding word. Modal particles are usually proclitic on the verb, as in Standard Modern Greek, because they occur preverbally and are stressless. However when followed by a preverbal clitic pronoun, they receive stress and become the stressed phonological host of the proclitic (Janse 2000). The example in (53), illustrates both stressless and stressed *na*.

(53) na péyo ná = to féro (Janse 2000)
CP I-go CP acc.3.sg I.bring
'I will go and bring her'

If Dawkin's (1916) data is correct, the above example could be represented phonetically as [na´peγo ´nato ´fero]. The ability for proclitics to attach outside the verbal domain clearly shows that preverbal clitics do not form a morphological unit with the verb.

From the perspective of grammaticalisation theory (Hopper&Traugott 1993, Heine, Claudi, Huennemeyer & Friederike 1991), Janse (2000) takes the position that clitic pronouns in Cappadocian are 'hybrid' (Heine, Claudi, Huennemeyer & Friederike 1991) falling somewhere between autonomous words and affixes. No analysis however is provided and no claims are made about the exact theoretical status of these clitics.

Based on Janse's data about enclitics and the general behaviour of clusters, it seems that there is strong evidence for treating these clitics as verbal affixes. Perhaps one way of accounting for the 'intermediate' behaviour of these clitics would be to adopt an analysis in which they do not attach uniformly. I would therefore argue that enclitics are morphologically attached to the verb, while proclitics exhibit properties of phrasal affixes. Using this analysis, it would be possible to explain

why enclitic and proclitic clusters are exactly identical, even though the former interact morphophonologically with the verb while the latter are phonologically enclitic to an adverb (even though syntactically proclitic to the verb). In addition, treating proclitics as phrasal affixes would explain why they can exhibit a mismatch between the syntactic host and the phonological host. In fact, as pointed out early on by Klavans (1980), the phrasal placement of affixes assigns them the ability to attach phonologically outside the domain within which they are syntactically placed. This ability is not observed in contemporary EP, where proclitics are phonologically and syntactically enclitic on the phrasal host[65].

b) Romanian (Popescu 2000, Monachesi 1998, 2000, Gerlach 2001a, Legendre 2000a)

One further example of mixed placement comes from Romanian. As in most Romance and Slavic languages, pronominal clitics in this language combine into clusters and occur immediately adjacent to the verb. Both in preverbal and postverbal position, they undergo morphophonological alternations and cannot be separated from the verb. Inflectional studies (Monachesi 1998, 2000) argue that clitic pronouns are uniformly word-level affixes. Intriguingly however, preverbal clitics can be phonologically enclitic on the preceding stressed word (regardless of its category):

(54) [ma.ma.l]$_{Pwd}$ place
mother-acc.3sg like.3sg
'mother likes him' (Popescu 2000)

(55) [nu-l] $_{Pwd}$ da
neg-acc.cl gives
'He does not give it' (Monachesi 2000)

---

[65] Condoravdi&Kiparsky (2001) provide a syntactic analysis for Cappadocian clitic placement and treats clitic pronouns as X^{max} clitics, i.e., word level units syntactically adjoined to a maximal phrasal projection. The problem with this view is that it leaves the purely affixal behaviour of postverbal clitics unaccounted for.

The fact that preverbal clitics have a phonological host outside their syntactic domain suggests to Gerlach (2001a) and Popescu (2000) that clitic pronouns in Romanian cannot be regarded as affixes.

Given the cohering relation between the verb and postverbal clitics and the affixal properties of clitic clusters, Monachesi (2000) argues that despite the behaviour of the clitic in (55), clitics in this language are realised as suffixes and morphologically part of the preceding host. The problem with this view is that it entails that preverbal clitics can be attached to any of the categories preceding them (e.g., nouns, adjectives, complementisers, etc.). Affixes however do not generally attach to more than one category. Legendre (2000a), in line with Anderson (in press, 1996), treats Romanian clitics as phrasal affixes. Phrasal affixation would be able to capture the promiscuous attachment of the preverbal clitics. But, as argued before for EP clitics, postverbal clitics in this language behave typically like verbal suffixes (cf. Chapter 3 and 8 for Anderson's phrasal affix approach to EP).

It seems then that the uniform attachment within morphology (as in Monachesi 2000) or within syntax (as in Legendre 2000a) raises problems because the data behaves neither purely like word-level affixes nor purely like phrasal affixes. Therefore one way of capturing the ambiguous behaviour of clitics in this language would be to attach postverbal clitics morphologically to a verbal stem and to attach preverbal clitics to a phrasal node.

c) Udi (Harris 2000, Crysmann 2002)

The Caucasian language Udi has recently attracted the attention of linguists because of the puzzling behaviour of its subject agreement markers (Harris 2000, 2002; Crysmann 2002). For reasons of space I will only summarise the aspects of the data that are more relevant for the discussion, however a very detailed study can be found in Harris (2002).

With verb forms marked Future, Subjunctive and Imperative forms, subject agreement markers are enclitic on the entire verb form (56), occur between a light verb and its incorporated element

such as a noun, adjective, or other element (57), or occur immediately before the final segment of the verb stem in simplex (monomorphemic) verbal stems (58).

(56) aq'-al-zu [receive-FUT-1SG] 'I will receive'

(57) k'al-zu-x-a [call-1SG-LV-PRES] 'I call'

(58) u-z-k-o [eat_1-1SG-eat_2-FUT] 'I will eat'

These agreement markers are allomorphic variants that are selected by the phonological and morphological properties of the verb (Harris 2000, 2002). One of these sets of allomorphs is selected by a small set of verbs, called inversion verbs (e.g., *-zu* and *–z* in the general paradigm become *–za* in the inversion paradigm; *-nu* and *–n* become *–va*, etc.).

What is intriguing is that in clauses with a focused constituent, the subject pronoun may be enclitic to the word in focus, as in (59), or to the phrase in focus, as in (60):

(59) zu nard -zu acipe
I.ABSL backgammon -1SG played
'I played BACKGAMMON' 'It was backgammon that I played'

(60) me isq'ar-mux mano aiz-i-q'un karx-esa?
this man-PL-ABSL which village-DAT-3PL live-PRES
'WHICH VILLAGE do these men live in?'

The subject pronoun can also attach to negative particle or to wh-pronouns, and the constituents carrying the subject pronoun can be separated from the verb by intervening words.

In addition, when the person marker is attached to the focused constituents it can have wide scope over coordination:

(61) a. evaxt'-t'u cax-exa cur-ax, evaxt'-t'u tast'a k'ok'oc-a k'ac?
when-2sg milk-say.pres cow.dat, when-2sg give.pres chicken-dat feed.absl
When do you milk the cows, and when do you feed the chickens?'

b. evaxt'-t'u cax-exa cur-ax da k'ok'oc-a k'ac tast'a?
when-2sg milk-say.pres cow.dat and chicken-dat feed.absl give.pres
'When do you milk the cows and feed the chickens?'

Interestingly, as pointed out by Crysmann (2002), even when the agreement marker attaches outside the verb its shape is still determined by the lexical properties of the verb.

(62) a. ma-n as-besa?
where-2sg work-do
'Where do you work?'
b. un ek'a-va aba?
you.erg what.abs-2sg know
'What do you know?'

What we have then is a person marker which, by default, realises verbal features on the verb (56-58) and which, under specific syntactic conditions, can also appear as a focus marker on not necessarily contiguous words (59-62). Particularly interesting is also the fact that the person marker can have wide scope when it is not attached to the verb. The data then is reminiscent of the EP clitic system in which proclitics can optionally have wide scope, but not enclitics.

Harris argues that the person marker constitutes an X° unit (clitics in the sense of Zwicky 1977) which is positioned as an endoclitic whenever it occurs word-internally, as in (57-58), thus violating the principle of lexical integrity (Harris 2000). Crysmann (2002) argues otherwise and explores a linearisation-based analysis within HPSG which treats subject agreement markers as affixes. Indeed, several reasons seem to suggests that treating Udi markers as affixes is more insightful. For reasons of space, I will not illustrate them but they include, among other aspects, the morphosyntactic information of the person marker, the alternation between agreement allomorphs, and the verb-internal appearance of the markers. From this viewpoint, one way of accounting for the behaviour of

Udi agreement markers would be to assume that they attach like genuine suffixes to the verb but that they constitute phrasal affixes when they appear in focused structures.

c) Synthesis

This section explored data from Cappadocian, Romanian and Udi, which showed that a given affix or affixal sequence can select both a morphological and a phrasal host. In Chapter 8, examples from Mari, a Finno-Ugric language, also show that recently grammaticalised affixes seem to have variable placement. The examples are repeated here for convenience and they show that the case markers can behave like a nominal suffix (63) or as a phrasal affix taking wide scope over a conjoined noun phrase (64).

(63) puškuδê-lan δa joltaš-lan (Luutonen 1997)
neighbour-DAT and friend-DAT
"to the neighbour and friend"

(64) puškuδo δa joltaš-lan
neighbour and friend-DAT

The PFM-analysis developed for EP cliticisation can straightforwardly account for the observed placement patterns. For Cappadocian and Romanian, for example, we can assume that realisation rules generate the clitic clusters and that placement rules determine their attachment: attachment would be morphological for enclisis and phrasal for proclisis. As to Udi, if our generalisation about 'mixed' placement in this language is on the right track, then we might assume that realisation rules generate the agreement marker as an affix and that placement rules determine the domain of grammar within which it attaches. I will leave a more refined treatment of this data for further research.

9.5 Summary

At the center of this chapter have been both the external and internal behaviour of clitic clusters. Externally, I examined dual placement and asymmetric attachment; internally, I investigated the invariant order of clitics inside the cluster. The leading idea of the analysis is that sequences of affixal clitics are derived as composed inflectional units through function composition. Whether the composed sequence of clitics is placed in preverbal or postverbal position is determined by an alignment function which defines a) direction of attachment and b) domain within which attachment takes place.

Placement is therefore stated only once, for the whole clitic sequence. Empirically, there are strong reasons for preferring this derivation and treatment of clusters. First, it makes it more straightforward to explain why clitic order remains invariant; second, it also enables us to predict that enclitic and proclitic clusters exhibit the same co-occurrence restrictions and the same morphophonology[66]. In addition, it also seems to predict that clitic sequences cannot be broken up be ordinary affixes, and that any changes in affix order affect the whole cluster not just parts of it[67]. without ever leaving clitics behind. So, in EP mesoclisis it is the whole clitic sequence that can appear in verb-internal position, not just part of it.

[66] This generalisation is true for most of the standard Romance languages, with very few exceptions (e.g., French in Miller&Sag 1997).

[67] Modern Greek however is an exception because in this language, clitic clusters can be broken up. Split clitics however are quite marked within cliticisation.

Chapter 10 On proclisis

This chapter addresses the morphology-syntax interface giving primary attention to the contexts triggering preverbal placement. As briefly alluded to before, EP proclisis is triggered by the presence of certain preverbal triggers, such as the negation marker, aspectual adverbs, wh-phrases, focused phrases and subordinating complementisers (Martins 1994, Duarte&Matos 2000, Crysmann 1997). The main goal will be to examine how these contexts can be placed in correspondence with the morphology without weakening the inflectional analysis proposed in the previous chapter. The claim underlying this goal departs substantially form the claims made by Vigário (1999b), Gerlach (2001a) and Crysmann (1999, 2002) who argue that an inflectional analysis of EP cliticisation is untenable on the grounds that proclitic triggers are not readily available to the morphology.

Formally, the interaction between the morphology and the proclitic contexts will be captured within the architecture of Lexical-Functional Grammar (Kaplan&Bresnan 1982, Bresnan 2001), using insights from Luís&Sadler (2003). Section 10.1 offers an empirical survey of proclitic contexts and argues with Anderson (2000), Luís (2003a) and Luís&Spencer (in press) that the affixal status of clitic pronouns is not affected by the syntactic nature of the factors determining clitic placement. Section 10.2 then introduces the theory of Lexical-Functional Grammar and explores the interaction of PFM with c-structure and f-structure syntax.

10.1 Preliminaries

The aim of the first part of this chapter is to argue, based both on empirical evidence and on previous studies, that the nature of the factors determining the enclitic-proclitic alternation does not invalidate the claim that EP clitics are generated as affixes.

Some of the preverbal elements triggering proclisis shall be surveyed and Anderson's (2000) account of EP clitic placement will be briefly discussed. Emphasis is given to Anderson's assumption that EP clitics constitute affixes like their Romance counterparts despite the factors triggering placement. It is shown that the PFM analysis of EP cliticisation proposed in Chapter 9 shares with Anderson's the insight that syntactic contexts can trigger morphological alternations (Luís 2003a, Luís&Spencer in press).

10.1.1 Inflectional views

Preverbal and postverbal placement of clitic pronouns in most Romance languages (e.g., Spanish, French, Italian) is conditioned by the finiteness features of the verb (Auger 1994, Miller&Sag 1997, Monachesi 1999). In Italian and Spanish, enclitics follow non-finite verbs and proclisis precede finite ones:

(1) a. Vederla ogni giorno (It.)
To see her every day.
b. La vedo ogni giorno
I see her every day

(2) a. Voy a dártela. (Sp.)
b. Te la voy a dar.
I am giving it to you.

As alluded to before, however, object pronouns in European Portuguese are not sensitive to the tensed or untensed properties of the verbal host. As illustrated below, we find enclitics and proclitics in both finite and non-finite clauses:

(3) a. Deu -lhe o anel.
gave -3sg.dat the ring
"(s/he) gave him/her the ring."

b. Prestar -lhes ajuda, seria possível.
to offer -3pl.dat help, would-be possible
'To help them, would be possible

(4) a. Não lhe oferecemos o anel.
Not 3pl.dat gave the ring
"(we) didn't give them the ring."

b. Para lhe perguntar, precisamos de saber onde mora.
to 3sg.dat ask need to know where lives
To ask him/her, we need to know where s/he lives.

In contrast to their Romance counterparts, EP clitic placement is sensitive to a specific set of words and phrases in preverbal position (Martins 1994, Duarte et al. 1995, Duarte&Matos 2000, Crysmann 1999, 2002). These elements, which are known in the Principles and Parameters/Minimalist literature as 'operators', include for example the sentential negation in (5) shows (compare the enclitic in the first clause with the proclitic in the second).

(5) As professoras deram -lhes livros, mas não lhes deram revistas.
the teachers gave 3pl.dat books; but not 3pl.dat gave magazines
´the teachers gave them books, but they didn't give them magazines'

Proclisis also takes place in a variety of subordinating clauses. Compare, for example, the alternation between enclisis and proclisis in (6).

(6) O Pedro encontrou -os, porque os procurou.
The Pedro found 3pl.masc.acc, because 3pl.masc.acc searched
'Pedro found them, because he searched for them'

In addition, a wide number of adverbial particles also trigger proclisis (Martins 1994). Examples with *também* 'also', *até* 'even' and *já* 'already', respectively, are given below:

(7) a. As crianças também o viram.
the children also 1sg.masc.acc see
'the children saw him, too'
b. *Já* te disse isso várias vezes .
already 2sg.dat said that several times
'I have already told you that several times'

Wh-phrases, either in main or embedded clauses, also provide the syntactic contexts for proclisis:

(8) a. Quantos livros nos compraste?
how-many books 2pl.dat buy
'how many books did you buy us?'
b. Eles querem saber quando te vais embora.
they want know when 2sg.refl go away
'they would like to know when you are leaving'

Likewise preverbal focus, but not topic, attracts the clitic into proclitic position (Martins 1994, Duarte et al 1995). In (9a) there is a topic NP, hence proclisis cannot occur.

(9) a. Deste livro, lembro-me mal.
of-this book, remember 1sg.ref vaguely
'this book, I remember it vaguely'
b. Deste livro me lembro muito bem.
of-this book 1sg.refl remember very well

This BOOK I remember very well.

These are some of the contexts within which proclisis occurs. A more detailed list of contexts is given by Duarte et al. 1995, Duarte&Matos 2000, Crysmann 2002.

Different accounts exist in the literature about how to best capture the factors triggering proclisis. With few exceptions, such as Anderson (2000) and Duarte&Matos (2000), the discussion of proclitic contexts is generally not interlinked with the affixal status of clitics. In particular, the fact that each one of the above proclitic triggers is syntactically visible - they constitute either single words (e.g., 5-7) or whole phrases (e.g., 8-9) - has been used to weaken the inflectional status of EP clitic pronouns (Gerlach 2001a, Crysmann 2002, Vigário 1999b). This claim has been explicitly formulated in Vigário (1999b) who argues that clitic pronouns must be regarded as word units because "the syntagmatic information relevant for the distribution of pronominal clitics" is not available in the morphology (1999b:223). Similar arguments have also been provided by Gerlach (2001a) who treats clitics as special word-level units (Gerlach 2001a) and by Crysmann (2002) whose analysis of clitics as 'morphosyntactic hybrids' is also motivated by the nature of the proclisis/enclisis alternation. For Anderson (2000), as shall be seen next, the fact that the proclisis-enclisis alternation is not dependent on finiteness properties of the verb does not constitute an argument against generating clitic pronouns as affixes.

10.1.1.1 Earlier treatment

As alluded to before, the theory of cliticisation developed by Anderson (1992, 1995, 1996, 2000) assumes that all affixal clitics are instances of phrasal affixation. Anderson's claim that Romance clitics are generated in the morphology as inflectional affixes is also adopted in this thesis[68]. Let us

[68] At various points of my argumentation I have shown that I do not share this view (cf. Chapters 3 and 8). In particular, I depart from Anderson by assuming that affixal clitics can also attach in the morphology (cf. also Spencer 2000, Monachesi 1999, Miller&Sag 1997). So, even though Anderson

now examine how Anderson's formalises the effect of proclitic contexts within his inflectional OT analysis of clitic placement.

In Anderson (2000) it is assumed - following insights from Barbosa (1996) - that the clitic patterns oberved above (cf.1-9) are motivated by a restriction which prevents pronominal clitics in EP from appearing clause-initially. This generalisation, which can be attested productively in Old French and Bulgarian (Halpern 1996), is generally referred to as the Tobler-Mussafia law. For EP, it is claimed that the placement of clitic pronouns is derived through the interaction between two constraints: one LeftMost constraint, such as (Edge (cl, L,Vº)), which requires the proclitic to be adjacent to the left edge of the verbal node; and a Non-Initial constraint, namely (Non-Initial (cl, IP or CP), which dominates the former by preventing the proclitic from being clause-initial. This should capture that EP clitics follow the verb only if the verb is initial within IP (or in some cases CP). Otherwise clitics precede the main verb.

In effect, Portuguese is exceptional within other Romance languages such as Italian or the Spanish in that proclitics cannot be initial in the clause. It is however not clear that this pattern is the result of the Tobler-Mussafia law, as suggested by Anderson (2000) and Barbosa (1996). The first problem with the assumption that placement in EP is sensitive to the Tobler-Mussafia Law is that in EP enclisis occurs even when the verb is not clause-initial, neither CP nor IP initial, as shown in (10). By the Tobler-Mussafia law, however, the clitic should be preverbal in the contexts.

(10) a. Dei -lhe o livro.
gave -3sg.dat the book
'(I) gave him/her the book'

b. O Pedro deu -lhe o livro.
the P. gave -3sg.dat the book

---

lays the foundations for an inflectional analysis of EP cliticisation, there are several aspects about his treatment that are not adopted in this thesis. For example, the idea that Romance clitics are phrasal affixes (cf. Chapters 3 and 8) and his approach to cluster derivation (Chapter 9).

'Pedro gave him/her the book'

c. O Pedro e a Rita deram -lhe o livro.
the P. and the R. gave -3sg.dat the book
'Pedro and Rita gave him/her the book'

To account for the fact that postverbal placement is not entirely subject to the Tobler-Mussafia law, Anderson (in press) and Barbosa (1996) postulate that the referential subjects in (9b-c) cannot therefore not be regarded as part of the IP within which the Tobler-Mussafia law applies because they are adjoined to IP. Definite subjects are thus treated as dislocated constituents, leaving the verb in IP-initial position. Counter-arguments to the proposal of Barbosa (1996), adopted by Anderson (in press), have been presented by Costa (2000), based on Ambar (1992) and Duarte&Matos (2000), which show that the adjunction analysis of subjects is not tenable for EP and that definite subjects must be placed in Spec-IP. Following Costa (2000), then, I assume that there is no empirical evidence supporting the adjunction to IP of definite subjects.

In addition, it also fails to capture the default status of enclisis, given that the Tobler-Mussafia law derives preverbal placement as the more productive type of placement. In older stages of Portuguese, as reported by Martins (1994), proclisis seems to have been the default type of placement. However, in contemporary EP the frequent misplacement of enclisis in adult speech and the early acquisition of enclisis by children suggest that enclisis constitutes the unmarked placement (Duarte et al. 1995). There is then insufficient evidence supporting the claim that the Tobler-Mussafia law determines the alternation between enclisis and proclisisis

A different position is assumed in this thesis. Affixal clitics in EP are not dependent on the edge of the clause but on a set of proclitic contexts or constructions which contain specific words or phrases in preverbal position[69]. This view is also adopted in non-inflectional studies by Duarte&Matos (2000), Crysmann (2002) (cf. 10.2 for alternative analysis). However, putting the

[69] For arguments agianst the view that proclitics are subject to the Wackernagel Law, see Chapter 8.

shortcomings of Anderson's analysis aside, it is worth considering the idea that the effect of syntactic contexts on clitic placement does not affect the affixal status of EP clitic pronouns. For Anderson, the status of Romance clitic pronouns is independent of the principles conditioning placement in each one of the languages. Regardless of whether the factors triggering placement are morphological or syntactic, Anderson assumes that clitic pronouns in Romance are affixes. Within the context of the other Romance languages, this is an important point because it can accommodate the fact that EP shares with the other Romance languages numerous similarities which would otherwise be left unaccounted for. This positions will also be adopted in this thesis.

Unfortunately, the exact way in which the morphology interacts with the syntax is not explicitly treated in Anderson's OT-analysis. Perhaps as a consequence of the OT-framework, within which the modularity of grammar plays no substantial role. In this thesis, a more explicit analysis of the interplay between morphology and syntax will be explored, within the architecture of Lexical Functional Grammar. It will be assumed that within this theoretical framework, PFM can interact with other components of grammar as an autonomous and parallel level of grammar (cf. 10.2).

10.1.1.2 PFM-proposal

Before placing the morphology in correspondence with the syntax, let us first brifly recapitulate the morphological analysis of clitic placement proposed in Chapter 9.

The PFM-analysis of proclisis and enclisis presented in the previous chapter is based on the assumption that there is one set of constraints deriving the affixal cluster and that there is another set of constraints defining whether the affixal cluster appears in preverbal or postverbal position. It has been suggested that proclisis placement in EP rules attach a given affixal cluster to a phrasal host within VP; enclitics, on the contrary, attach to a morphological verbal host and constitute therefore morphological affixes (cf. Chapter 8 for arguments for asymmetry). The crucial aspect however is that clitics are generated as affixes.

As to the factors triggering clitic placement, finiteness does not play any role in EP unlike in other Romance languages. To compensate for the lack of a morphological property, a formal feature has been introduced in the morphology to determine the alternation between the preverbal or postverbal placement of affixal clitics. In Chapter 9, therefore, it has been assumed that clitic placement is mediated at the morphological level by a special, purely formal, feature [Restricted:Yes/No]. This means that verb forms carrying the feature [Restricted:yes] are associated with proclitic placement while the default feature [Restricted:no] triggers enclisis.

How this formal feature interacts with placement, within the morphology, is illustrated below, using exampes from Chapter 9. In PFM, a paradigm function is a function which applies to the stem of a lexeme to derive a cell in that lexeme's paradigm. Paradigm functions then generate morphological cells. In (11), a Paradigm-Function is given which defines how cliticised verbs in EP are derived (the whole derivation of each cliticised verb can be consulted in Chapter 9). The format of the PF has been extensively discussed in Chapters 6 and 9.

By (11b), the affixal cluster is derived in two stages: (11b-i) generates the cluster and (11b-ii) determines placement, with the feature [Restricted:yes] triggering proclisis (for placement in other Romance languages, cf. 9.3.3.2 in Chapter 9)

(11) Paradigm-Function for cliticised verbs in EP

PF (<X, (σ_{1+2}>)

$=_{def}$ { a. base { i. stem X-
ii. verbal suffixes RR_{II} (RR_I (σ_1)
b. cluster { i. exponence ($RR_A \circ RR_B \circ RR_C$) ($\sigma_2$)
ii. placement { < <aff,aff >, V° >, if Vf [Restricted: yes]
< stem, <aff,aff > }, elsewhere

It is worth pointing out that the 'restricted' feature is just one of a number of technical devices one might propose in order to determine between enclitic and proclitic placement. This feature has two welcome consequences. It enables us to capture that there are two types of conjugations, the enclitic and the proclitic ones. What we are effectively saying is that this formal morphological feature divides verb forms into two conjugations (or subparadigms) - restricted and neutral – which don't themselves directly express any feature value, very much like the 'restricted' verbal paradigm in Somali which is selected by phrasal properties of the clause, but which in itself carries no specific feature value (see below).

In addition, note that the feature 'restricted' used in the analysis directly captures the claim that proclisis constitutes the marked position for clitic clusters which is solely found with verb forms associated with the [Restricted:yes] feature[70]. Under all other circumstances we obtain enclisis, which is regarded as the default, under the assumption that [Restricted:no] is the default feature value. Proclisis is regarded as marked morphology. The analysis is supported by data from language acquisition and adult speech which shows that enclisis constitutes the unmarked (and more productive) case in EP, unlike in most of the other Romance languages (Duarte et al 1995).

10.1.2 Summary

This section surveyed the contexts within which EP proclisis occurs and argued these contexts should not be used to invalidate the inflectional status of clitic pronouns. Based on this assumption, the Paradigm-Function treatment of clitic placement in Chapter 9 introduces a formal feature which makes it possible to capture the idea that clitic pronouns constitute affixes, even though the alternation between enclisis and proclisis is triggered by syntactic information. A similar assumption about the separation between affixal status and placement triggers is adopted by Anderson (2000).

Even though purely formal features should be avoided on the grounds that they are not meaningful, it is important to observe that the existence of such formal features are not atypical in inflectional morphology. For example, information as to whether verbs belong to the first, second or third conjugation class in Romance cannot be associated to any meaning, although it is crucial for determining which tense and agreement suffixes combine with a given lexeme.

In addition, there are ordinary inflectional systems which display morphological alternations that seem to be triggered by clausal properties. Inside the morphology, at least, the alternation appears to have no effect on the meaning of the verb forms.

For example, in Somali there is a morphological distinction between restricted conjugation and normal conjugation. Comparing both paradigms with each other we observe that the restricted paradigm is characterised by a reduction of some agreement features of the normal paradigm. There are different forms in the restricted paradigm which do not appear in the normal paradigm.

(11) Normal and Restricted Paradigm of the past tense of the prefix verb *imaansho* 'to come' (Slovacchia, M. et al. 1995)

	Normal Paradigm	Restricted Paradigm
1	imid	imid
2	timid	yimid
3	yimid	yimid
3f	timid	timid
1p	nimid	nimid
2p	timaadeen	yimid
3p	yimaadeen	yimid

The intriguing property is that the shape differences between one conjunctiona and the other do not signal any semantic distinction. Instead, the restricted forms are semantically exactly identical to the corresponding normal forms. As pointed out by Lecarme (1991), based on Andrzejewski (1956), the 2nd and 3rd plural verb forms on the normal paradigm express precisely the same features as the

[70] In our treatment of EP, the feature 'restricted' is adopted from the restricted conjugation class in

corresponding verb forms in the restricted paradigm. Instead, the selection of the restricted conjugation is dependent on whether the subject of the clause is focused or not (Slovacchia, M. et al. 1995, Lecarme 1991). In fact, subject focusing in Somali entails a number of consequences, one is that the focused subject NP loses its subject mark (12b), the other – and more relevant for the present discussion - is that the agreeing verb has a special type of conjugation, the restrictive paradigm (Re), characterised essentially by the reduction of some agreement features of the normal paradigm (12c). (FM = focus particle)

(12) a. niman-kaas baa hilib cunayá (Slovacchia, M. et al. 1995)
men-those FM meat eating (Re)
'THOSE MEN are eating meat'
b. * niman-kaasu baa hilib cunayá
men-those FM meat eating (Re)
c. * niman-kaas baa hilib cunayaan
men-those FM meat eating (N)

Assuming that the choice of the 'restricted' conjugation in Somali is dependent on syntactic focus, then we may conclude that the feature 'restricted' in Somali can be effectively regarded just as a formal feature whose role it is to signal a specific type of clause. The data then supports the idea that morphological alternations can be purely the reflex of the properties of the clause. For this type of phenomenon, formal features are needed within inflectional morphology to capture the fact that the morphological alternation is simply a reflex of a specific type of syntactic context. An identical case seems to take place in Potowatomi (an Algonquian Language), where the choice between the

---

Somali (cf. 10.1.1).

conjunct and the indicative moods are dependent on whether the clause is subordinate or not (Hockett 1948)[71].

10.2 The morphology-syntax interaction

This section discusses EP clitic placement from the point of view of the interplay between morphology and syntax. The main goal will be to show how to place the formal feature [Restricted] (cf. 10.1.1.2), which resides in the morphology proper, in correspondence with the proclitic contexts. Formally, the morphology-syntax correspondence will be explored within the framework of Lexical-Functional Grammar (Bresnan&Kaplan 1982, Bresnan 2001).

Fundamental to the LFG architecture is the assumption that the modules of grammar consist of several different parallel structures[72]. Each of the different structures of LFG has a distinct formal character and models a different aspect of the structure of language. No structure is derived from another structure and the relationship between structures is defined by a mapping or correspondence function. The main modules of grammar are c(onstituent)-structure, which models phrase-structure relations, and f(unctional)-structure, which encodes grammatical relations. Within this parallel design of the LFG architecture, the morphology also constitutes an autonomous level, namely the level within which words are derived and assembled. Following previous work by Sadler&Spencer

---

[71] There is also evidence from other clitic systems suggesting that the morphology must interface with the syntax in order to capture placement alternations. In Greek varieties, such as the Cappadocian dialects, Janse (2000) and Luís (2003a) have argued that there is evidence suggesting that these clitics constitute affixes (cf. Chapter 9). Note that the factors determining preverbal and postverbal placement are similar to the ones found in EP. Also, the factors conditioning the placement of Udi agreement markers share similarities with those surveyed above for EP (Harris 2000).

(2001), Sadler&Nordlinger (2003), Luís&Sadler (2003), Spencer (2003), Otoguro (2003), the model of morphology I shall incorporate into LFG is that of Paradigm-Function Morphology (Stump 2001, Spencer ms) [73].

To account for proclitic placement it will be argued that by allowing the morphology to interact with other components of grammar, within LFG, it will be possible to assign proclitic contexts a syntactic property and invest them with the ability to select the relevant cliticised verb form. Before formalising the proposal, a brief introduction to the basic aspects of the morphology-syntax interface will be provided, based on current developments within Lexical-Functional Grammar by Sadler&Spencer (2001), Sadler&Nordlinger (2003) and Kaplan&Butt (2002). The interface will be discussed primarily from the point of view of the c-structure and f-structure representation of EP affixal clitics (10.2.1). Having clarified some of the issues, an account will be sketched of the effect of the proclitic contexts on clitic placement (10.2.2).

10.2.1 Affixal clitics in LFG

In Lexical-Functional Grammar, much recent work has explored the idea that words and affixes are different means of encoding the same grammatical relations (Bresnan 2000; Börjars, Vincent and Chapman 1996, Nordlinger 1998, Sadler&Spencer 2001, Sadler&Nordlinger 2003, Luís&Sadler 2003, Spencer 2003, Otoguro 2003). The separation in LFG of c(onstituent)-structure from f(unctional)-structure captures the possibility that syntactic and morphological elements – affixes, words and phrases - may express the same f-structure information in a natural way

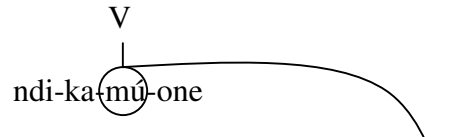
---

[72] For reasons of space, the survey of the general architecture of Lexical-Functional Grammar will be kept to a minimum. A detailed introduction to LFG is provided in Falk 2001. More advanced surveys are provided by Bresnan 2001 and Dalrymple 2001.

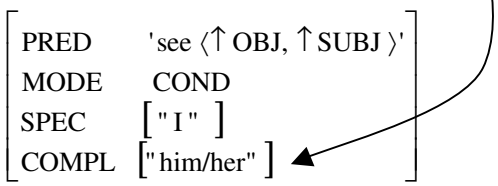
[73] Other levels of representation include prosodic-structure, argument-structure, semantic-structure and information-structure, but these will not be of concern here.

(Bresnan&Mchombo 1987). A striking example of the functional equivalence of affixes and phrases is found in the contrast between the English noun phrase *him* in 'I see him' (13) and the Chichewa object marker *mú-* in the verb form *ndikamúona* 'I see him/her' (14), in which the same grammatical information is realised by distinct expressions in c-structure (Nordlinger&Bresnan to appear):

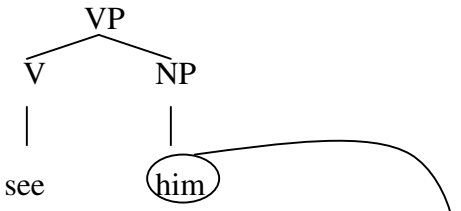
(13) a. c-structure (fragment)



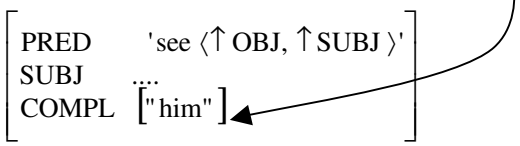
b. f-structure (fragment)



(14) a. c-structure (fragment)



b. f-structure (fragment)



(13b and 14b) illustrate the c-structure representation. C-structures then are phrase-structure trees which represent 'surface' precedence and dominance relations in a language, adopting standard X'-theory (e.g., Chomsky 1986). The main departure from standard X'-theory assumptions is heads are optional and used only if required. Unlike derivational theories, there is no movement and therefore words are assumed to be base-generated directly in their surface positions. One further principle that holds in c-structure is the Lexical Integrity Principle. A consequence of this principle is that only fully inflected words can be terminal elements of c-structure trees. Elements 'smaller' than a morphological word are disallowed as syntactic terminals (Bresnan 2001). The more universal f-structure encodes grammatical relations (e.g. SUBJect, OBJect), tense features (e.g, TENSE), nominal features (e.g, CASE , NUMber, GENder) and the predicate attribute PRED. As illustrated in

(13a) and (14a), f-structures are attribute-value matrices. The contents of the f-structure come from the lexical items of the sentences or from annotations on the nodes of the c-structure (cf. Bresnan (2000) for detailed discussion about the relationship between f-s and c-s).

One important consequence of enabling parts of words to contribute the same f-structure information as phrases, as illustrated by (13), is that elements realising the argument function of the predicate are not tied to predefined argument position in phrase-structure, as assumed in derivational frameworks of syntax, such as Government & Binding Theory (Chomsky 1981), Principles and Parameters Theory (Chomsky&Lasnik 1993) or the Minimalist Programme (Chomsky 1995).

In the next sections, I illustrate the f-structure and c-structure representation of EP affixal clitics (10.2.1.1-10.2.1.2), giving emphasis to the morphology-syntax mapping.

10.2.1.1 F-structure description

As alluded to above, parts of words also contribute information to the f-structure. In (13) above the information OBJ in the f-structure comes from the Chichewa prefix 'mú'. The ability of affixes to contribute f-structure descriptions is captured in classical LFG by associating affixes directly with f-structure descriptions. Within a morphemic approach to morphology, affixes are represented as sublexical entries which carry f-s information.

However, under a realisational approach to morphology (cf. Chapter 2), f-structure descriptions cannot be associated directly, because affixes are not regarded as meaning-form pairs (Kaplan&Butt 2002, Sadler&Nordlinger 2003). In the morphology proper, affixes (or morphological exponence) constitute the (phonological) realisation of bundles of m-features. In Chapter 9, we have been

concerned with the morphology-internal aspects of the data.The following shows some sets of m-features, together with their exponents, as proposed in Chapter 9 within PFM[74]:

(15)

	m-form	m-features
a.	*lhes*	{Case: Dat, Per: 3, Num: Pl}
b.	*os*	{Case: Acc, Per: 3, Num: Pl, Gen: Masc}.

In (15a) the affixal clitic *lhes* realizes the set of m-features {Case: Dat, Per: 3, Num: Pl} and in (15b) the affixal *os* constitutes the exponent for the m-features {Case:Acc, Per: 3, Num: Pl, Gen: Masc}.

Based on the morphological representations of affixal clitics assumed in Chapter 9, we now need to associate m-features with syntactic information. Following Sadler&Spencer (2001), Kaplan&Butt (2002), Luís&Sadler (2003), Sadler&Nordlinger (2003), an interface between morphology and syntax shall be assumed which places morphological features in correspondence with syntactic information[75]. The morphology-syntax interface can be realised by a mapping from m-feature to f-feature. This idea is illustrated in (16), where the list of m-features is placed in correspondence with the relevant f-structure description.

---

[74] These examples are presented in tabular form purely for convenience. We should stress that PFM consists of ordered sets of realisation rules and there is nothing that corresponds to lexical entries for individual affixes.

[75] In many cases, the mapping is straightforward, and (↑TNS) = PAST) in f-structure corresponds to [Tense: Past] in the morphology. However, more complex correspondences exist which indicate that the direct association of f-structure descriptions (i.e., a morphemic approach to morphology) is problematic both at the level of pure morphology (cf. Stump 2001, Spencer in press, cf. also Chapter 2, for form-meaning mismatches) and at the level of the morphology-syntax interface (Sadler&Nordlinger 2003:13, on case stacking). In the case of EP affixal clitics, we adopt the morphology-syntax mapping on the grounds of theoretical consistency only.

(16)

m-features		f-structure descriptions
Case: Acc	⇔	(↑OBJ PRED) = PRO
Person: 3	⇔	(↑OBJ PER) = 3
Number: Sg	⇔	(↑OBJ NUM) = SG
Gender: Masc	⇔	(↑OBJ GEN) = MASC

F-structure descriptions in (16) combine to provide the full f-structure description associated with a given affix.

(17) {Case: Acc, Per:3, Num:Sg, Gen: Masc} ⇔ (↑OBJ PRED) = PRO
(↑OBJ PER) = 3
(↑OBJ NUM) = SG
(↑OBJ GEN) = MASC

10.2.1.2 C-structure

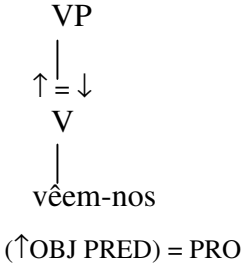
The above mapping correctly treats preverbal and postverbal clitics as affixes that are syntactically exactly identical by associating them to the same f-s information. In c-structure, however, enclitics and proclitics differ with respect to the phrasal node they are associated with: the syntactic (or f-structure) information is associated with the verb form, for enclitics, and it is associated freely with either the V or the VP node in the case of proclitics. That is, we propose that the placement rules themselves (cf. above 11b-ii) are crucial in determining which elements the syntactic (f-structure) information is associated with.

a) Enclitic and proclitic affixes

As Chapter 9 argued, enclitics are attached at the level of the Vstem. They form with the verbal stem a typical morphological cell. In c-structure, therefore, the verb-enclitic unit is syntactically opaque and appears under one single c-structure node as in (18a). The syntactic information associated with the clitic (represented here for convenience simply as (↑OBJ PRED) = PRO) is associated with the

whole word form in the annotated c-s (18a). The f-structure representation for (18a) is given in (18b)[76].

(18) a. c-structure fragment



b. partial f-structure

$$\begin{bmatrix} \text{PRED} & \text{'see } \langle \uparrow \text{OBJ } \uparrow \text{SUBJ} \rangle \text{'} \\ \text{SUBJ} & .... \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{PRO} \\ \text{PERS} & 1 \\ \text{NUM} & \text{PL} \end{bmatrix} \\ \ldots & \end{bmatrix}$$

Proclitics also form a cell with the verb, but they are not part of the verbform; the cell containing the proclitic and the verb is 'discontinuous' in the sense that the proclitic is not attached to the verb. This idea can be represented as in (19).

(19) a. verb-enclitic cell: <damos-lhe>

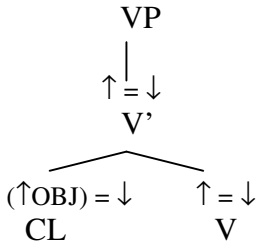
b. proclitic-verb cell:<lhe, < damos>

Different proposals have been made within LFG to account for clitic structures which represent clitics as syntactically independent units. Grimshaw's (1982) early analysis for French treats them as syntactic CL categories as in (20a). Sadler (1997) treats clitics in Welsh as non-projecting words (20b). These studies however represents clitics as syntactic units.

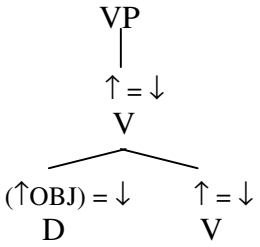
---

[76] A similar analysis would also be assumed for pronominal clitics in French (Auger 1994, Miller&Sag 1997), Spanish (Andrews 1990) and Italian (Monachesi 1999).

(20) a.

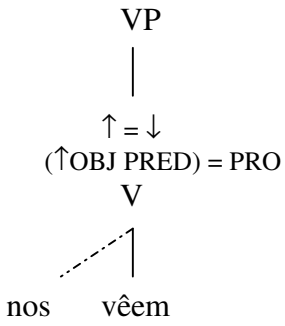


b.



Given that proclitics are neither morphologically part of a stem nor syntactically independent units, this study takes the position that a somewhat more complicated interface between the morphology and the (external, structural) syntax is needed than that commonly assumed (Luis&Sadler 2003). We therefore propose an analysis which introduces a modest extension to LFG syntax. We propose that proclitics constitute affixes without c-structure representation (21a) and associate pronominal f-structure information with V and VP phrasal nodes (21a-b).

(21) a. c-structure fragment



b. partial f-structure

$$\begin{bmatrix} \text{PRED} & \text{'see } \langle \uparrow \text{OBJ } \uparrow \text{SUBJ} \rangle\text{'} \\ \text{SUBJ} & \text{....} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{PRO} \\ \text{PERS} & 1 \\ \text{NUM} & \text{PL} \end{bmatrix} \\ \text{...} & \end{bmatrix}$$

It is crucial to observe that the proclitic in (21a) is not positioned under a syntactic node. The dashed line in (21a) signals phrasal affixation. As to the f-annotation of the proclitic, it can be associated freely with V or VP. The V is chosen when the proclitic has narrow scope (either because it occurs on its own in non-coordinated clauses (21a) or because it is repeated on each conjunct in coordinated structures); it is associated with the VP when the proclitic has wide scope[77].

b) Phrasal affixation and lexical integrity

What are the consequences of viewing EP pronominal clitics as phrasal prefixes? The most important consequence is that such an analysis implies a somewhat more complicated interface between the morphology and the (external, structural) syntax than that commonly assumed. In particular, we require the domains of morphology and syntax to overlap in a manner which is certainly challenging for the Principle of Lexical Integrity as normally understood.

An important aspect of this principle is the strict separation between morphology and syntax. Within this analysis, the building blocks of morphology are different from those of syntax, and structural syntax can only represent morphologically complete words. The principle of lexical integrity as a leading idea in LFG (and other lexicalist theories) has lead to many insightful and innovative analyses on a range of phenomena. It does, however, embody the simplifying assumption that the mapping between morphological and syntactic constructions is trivial.

As the data on EP proclitics indicates, the interaction of morphology and external syntax can be more complex. Phrasal affixation is a well established phenomenon proposed for a variety of clitics cross-linguistically (cf. Chapter 8). For EP proclitic, it is also the only adequate representation capable of capturing the intermediate status of EP proclitics. Given the placement rules provided for proclitics in Chapter 9, the domains of the morphology and the syntax must interact in such a way as to allow the clitic cluster (i.e. a sequence of affixes) to select a host in the syntax. In this study, an attempt has been made at representing phrasal affixes in c-structure without incurring a violation of

[77] Cf. Coordinated structures are examined in Luís&Sadler (2003).

the lexical integrity. In accordance with the Principle of Lexical Integrity proclitics are not represented under an independent syntactic node, and also the verb which forms a discontinuous cell with the proclitic comes out of the morphology as an independent word. It can therefore appear under a syntactically autonomous node[78].

Summing up, this section incorporated PFM into LFG and explored cliticisation within the wider context of the morphology-syntactic interaction. It has been shown that LFG can capture the fact that preverbal and postverbal clitics are syntactically exactly identical, by associating them to the same f-structure information. At the configurational level, on the other hand, a distinction has been made between the c-structure representation of morphological affixes and phrasal affixes. Based on the brief introduction to the morphosyntax of cliticisation presented above, this section will account for the effect of proclitic contexts on clitic placement by further exploring the relation between syntactic information and its morphological expression.

10.2.2 Proclitic placement

Finding a common denominator for the set of syntactic contexts has proved problematic. Syntactic analyses have generally tried to identify configurational similarities by moving most proclitic triggers to functional nodes within CP (Salvi 1990) but there appears to be little independent motivation for some of these assumptions; other linguists have accounted for the clitic patterns by placing proclitic triggers in different syntactic positions thus contributing towards a proliferation of functional projections (Martins 1994, Madeira 1993, Rouveret 1999). In addition, even though restricted subsets of these constructions do seem to constitute natural classes (e.g. downward

[78] More detailed implications of phrasal affixation for LFG, such as their representation in the phonological string, shall be left for further study (for insightful discussion of the role of, and problems with, this principle see Ackerman and Lesourd 1997 and Ackerman and Webelhuth 1998).

monotone quantifiers; see Crysmann (2002) for some insightful discussion of these cases), it is not plausible to try to extend these natural classes to encompass all proclitic triggers. In particular, discourse-informational effects have been argued to be the main force driving object pronoun placement in EP (McConvell 1996), but it is difficult to see how this claim could be extended to subordinating complementisers and conjunctions.

In our view, therefore, there is no common configurational or semantic/discourse explanation in the synchronic grammar for procliticisation. It shall be assumed that prefixal linearisation is a morphological reflex of an abstract syntactic feature, common to the set of syntactic construction types in (5-9). We suggest that the semantic core of this abstract syntax concerns the discourse notion 'Non-Neutral' and that the set of constructions is in the course of grammaticisation.

In Luís&Sadler (2003), this has been captured by positing an abstract syntactic feature (↑TYPE) = NON-NEUTRAL associated with the construction types which give rise to phrasal prefixation[79]. So for example, the f-structure for a negative clause such as (23a) contains the (abstract syntactic) information TYPE = NON-NEUTRAL in (23b):

(23) a. O João não me deu o livro.
the João not 1sg.dat gave the book
"João didn't give me the book."

b. f-structure

[79] The attribute TYPE is used in LFG to designate types of clauses, such as TYPE = Q for Wh-clauses (Kaplan&Bresnan 1982) or TYPE=V2 (Sells 2001). In Sells (2001) extensive motivation is provided to motivate the abstract syntactic feature TYPE=V2 which plays a role in selecting 'special' morphology in Swedish.

$$\begin{bmatrix} \text{PRED} & \text{'give } \langle \uparrow \text{OBJ}, \uparrow \text{OBJ2}, \uparrow \text{SUBJ} \rangle \text{'} \\ \text{SUBJ} & \text{....} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{PRO} \\ \text{PERS} & 1 \\ \text{NUM} & \text{PL} \end{bmatrix} \\ \text{POL} & \text{NEG} \\ \text{TYPE} & \text{NON - NEUTRAL} \\ \text{...} & \end{bmatrix}$$

By mapping the contexts onto TYPE = NON-NEUTRAL, we assume that this feature identifies the syntactic contexts that are responsible for proclisis. This feature stands for various types of triggers, some purely syntactic other semantic or discourse dependent. It is therefore assumed that these contexts belong to a disjoint list of syntactic constructions which contain one of the set of proclitic triggers in preverbal position. Only clauses in which the trigger appears preverbally will be associated with this abstract feature. The reason that post-verbal wh-phrases etc. don't trigger proclisis is because the TYPE NON-NEUTRAL attribute is only associated with preverbal elements. In other words the attribute TYPE NON-NEUTRAL is 'realized' so to speak by a combination of (a) wh-phrase, quantified phrase, negation, focus, etc. (b) preverbal c-structure position[80]. Thus, *in situ* wh-phrases (24a) or clauses with postverbal adverbs (24b) would not trigger enclitic placement.

(24) a. Compraste -nos quantos livros? (cf. 8a)
bought -2pl.dat how-many books
'you bought us how many books?'

b. Disse -te já várias vezes. (cf. 7b)
told 2sg.dat already several times
'I have already told you several times'

The question now is how to enable the TYPE feature to select phrasal prefixation for pronominal objects. The idea will be to place the formal feature [Restricted:yes], which resides in the

morphology, in correspondence with the proclitic constructions triggering proclisis, thus treating preverbal placement as a morphological reflex of specific syntactic information.

Recall that the formal feature [Restricted], which operates inside the morphology, determines the selection between preverbal and postverbal placement. Being morphology-internal, the feature [Restricted] is not relevant to the interface. For this reason preverbal and postverbal clitics are syntactically exactly identical, since they are associated to the same f-structure information. The [Restricted] feature will now be placed in correspondence with the proclitic constructions (in this sense, their functional structure). The idea is that the "special" morphology is selected by placing [Restricted] in correspondence with the abstract syntactic information TYPE = NON-NEUTRAL, perhaps by means of a mapping constraint along the following lines[81]:

This feature, which is needed for morphology-internal reasons, will now be placed in correspondence with the proclitic constructions (in this sense, their functional structure). The idea is that the "special" morphology is selected by placing the restricted feature, within the morphology proper, in correspondence with the abstract syntactic information TYPE = NON-NEUTRAL, perhaps by means of a mapping constraint along the following lines[82]:

(25) $(\mu(\uparrow) \text{Restricted}) = \text{Yes iff } (\uparrow\text{TYPE}) =_c \text{NON-NEUTRAL}$

This constraint states that if the f-structure of a given clause contains the syntactic information TYPE = NON-NEUTRAL, the cliticised verb form must be "restricted". In this analysis, then, preverbal placement is regarded as a case of marked morphology, i.e. the morphological expression

---

[80] I leave the formalization of this generalisation for further investigation (Luís&Otoguro in prep.).

[81] Luís&Sadler (2003) assume a correspondence between the placement function and the syntactic feature. Here I place the syntactic feature in correspondence with the formal morphological feature.

[82] Luís&Sadler (2003) assume a correspondence between the placement function and the syntactic feature. Here I place the syntactic feature in correspondence with the formal morphological feature.

of specific functional information. Also, by virtue of the 'iff', (25) states that the converse holds: if the verb form bears the feature value [Restricted: Yes], then TYPE = NON-NEUTRAL.

10.3 Summary

Incorporating PFM within LFG has made it possible to explore the interaction between the purely inflectional properties of cliticisation and the syntactic aspects of the data. Particular emphasis has been given to the phrase-structure status of phrasal affixes and the effect of proclitic contexts on clitic placement[83].

At the centre of the chapter has been the claim that the syntactic conditioning of proclisis does not invalidate an inflectional analysis of cliticisation. Given that the wide range of recent studies which argue the contrary (e.g., Gerlach 2001a, Crysmann 2002, Vigário 1999b), it seemed crucial to show that an account of the interaction between morphology and syntax is possible. The analysis of syntactic conditioning of proclisis however shows that the architecture of LFG permits a simple and intuitive analysis of the effect of syntax by assigning a common functional feature to proclitic constructions and by investing these contexts with the ability to select the inflectional proclitic-verb unit. The idea that inflectional morphology may sometimes be just a marker for specific syntactic contexts is supported by cross-linguistic evidence from Somali. In this language, it was shown that the restrictive and extensive conjugation paradigms are equivalent in all respects, except that the former is generally selected by clauses containing a subject focus (Lecarme 1991). An analysis along the lines of that proposed for EP might therfeore be adopted for this language.

---

[83] Other topics such as the scopal behaviour of proclitics and the distribution of clitics in auxiliary-verb structures have been previously treated in Luís&Sadler (2003) and Luís&Spencer (in press). The behaviour of clitics with respect to restructuring verbs shall be left for future study. However restructuring verbs will behave essentially in the same way that auxiliary verbs do provided there is

One further assumptions underlying this analysis is that two or more syntactic contexts may trigger the same morphological alternation without forming a natural class. This observation is not only true for proclitic triggers in EP, but also for unrelated languages, such as Udi (Harris 2000, 2002; Crysmann 2002), and Cappadocian (Janse 2000, Condoravdi&Kiparski 2001). In Chapter 9 it was suggested that these languages seem to exhibit an asymmetric placement of affixal units (i.e., clitic pronouns in the case of Cappadocian and subject agreement markers in the case of Udi). Intriguingly, similar syntactic constructions as found in EP determine whether these affixes are realised as phrasal affixes or as morphological affixes. To us, the fact that most of these contexts do not bear any clear relation with each other suggests that placement alternations constitute grammaticalised distinctions[84]. The account has shown what an LFG approach might look like and how the morphology can be placed in correspondence with the syntax through f-structure. The analysis however does not stop the grammar from relating other properties, such as quantifier types or prosody, to the TYPE NON-NEUTRAL attribute[85]. Note that I am assuming that neither semantics nor prosody can be used as sole explanation for preverbal triggering. Finally, as to the representation of cliticised verbs in phrase-structure, it has been argued that neither enclitics nor proclitics can be represented in c-structure as terminal elements, on the grounds that affixes are disallowed as syntactic terminals (Lapointe 1980, Bresnan 2001).

---

a relevant level at which the two verbs form a single complex predicate (Andrews and Manning 1999).

[84] Harris (2002) points out that the Udi facts result from an unusual set of historical changes. Each element is perfectly natural in itself, but the result of combining them in one language is unusual. One would therefore not expect the triggering set to form any kind of natural class.

[85] Semantic properties of preverbal triggers have been discussed in Crysmann (1999) and the effect of prosody on preverbal triggers by Duarte&Matos (2000).

Chapter 11 Conclusion

To conclude, I shall briefly summarise the proposals made in this thesis.

The Paradigm-Function analysis

Cliticisation in EP has been defined as an inflectional phenomenon, with clitic pronouns being treated as verbal affixes. The fact that there are robust signs of morphological construction between the enclitic and the verb indicated to us that enclitics to a verbal stem as inflectional suffixes. The term 'clitic' therefore has been used merely as a descriptive cover term, not as a theoretical concept. Similar claims have been made for other Romance languages (Monachesi 1999, 2000; Miller&Sag 1997, Brines 2000, Andrews 1990).

It has also been shown that EP clitic pronouns are distinct from clitic pronouns in the other Romance languages, in various ways. There is, on the one hand, the fact that enclitics can also appear verb-internally and, on the other, the fact that proclitics exhibit phrasal behaviour. To capture these properties, it is claimed that affixal clitics in EP attach asymmetrically: while enclitics are genuine verbal suffixes (which appear either as outer layers or as inflections on an infinitival stem), proclitics constitute phrasal prefixes. As verbal suffixes, they can appear either verb-finally or verb-internaly; as phrasal affixes, proclitics are generated in the morphology, but placed in the syntax (drawing on insights from Anderson 1992 and Spencer 2000). These insights have been incorporated within the theory of Paradigm-Function Morphology (Stump 2001), an 'inferential-realisational' theory of morphology. One important characteristic of this theory is that morphological forms are determined on the basis of the full feature content of the word form to be realised. The morphological formatives themselves are not lexical entries with their own feature sets or meanings.

Although the analysis of EP clitic system relies heavily on the theory of Paradigm- Function Morphology, a number of modifications have been introduced to the standard theory put forth in Stump (2001). The key differences are:

a) The paradigm function (PF) deriving cliticised verbs is defined in terms of two sets of morphological features, the set σ_1 - for tense, agreement and other ordinary verbal features - and σ_2 for clitic-related features such as case, person, person and gender, rather than in terms of just one set of features. This entails that a PF comprises two layers of affixation: one layer realising σ_1 features and another layer for the features σ_1. Generally, the sequence of affixes realising σ_1 appears as the inner layer of inflection, while the affixal clitics appear as the outer layer. The result of defining two 'layers' is that we can easily explain the interaction between inner and outer layers of inflection as found in 'mesoclisis'. Here we can nicely capture that each layer beahves as a whole unit with respect to mesoclisis.

b) The PF is defined over the stem of the lexeme, rather than the root. This makes it easier to capture the idea that there is a class of 'meaningless' stems in morphology, which constitute the input to exponents of morphosyntactic features (Spencer ms). In addition, it also facilitates the derivation of mesoclisis which, under the present proposal, is simply derived as affixation to an infinitival stem.

c) One crucial proposal has been the separation of the exponence (i.e., realisation) of clitic features from the placement (i.e., linearisation) of the exponents, so that we can generalise over each separately. The new format of the realisation rules deriving clitics is given in (1). Unlike standard realisation rules, these rules do not refer to the host which serves as input to the rule and therefore do not define whether affixes are suffixal or prefixal.

(1) a. $RR_{A1, \{Person:3, Refl:+\}, V}(\sigma) =_{def} [se, \sigma]$

b. $RR_{B5, \{Case: Dat, Person:3, Number: Sg\}, V}(\sigma) =_{def} [lhe, \sigma]$

c. $RR_{C1, \{Case: Acc, Person:3, Number: Sg, Gender: Masc\}, V}(\sigma) =_{def} [o, \sigma]$

d) If two or more realisation rules occur, they are composed into one sequence of exponents through function composition. So, if we combine the rules in (1a) and (1b) we obtain the affixal string (2).

(2) $RR_{C\{\}} \circ RR_{B5} \circ RR_{A1}(\sigma) =_{def}$ [<*se-lhe*>, σ]

This means that we no longer generate a clitic string by successive cycles of attachment of a single clitic. Rather, two or more realisation rules are combined as a composed function yielding a sequence of affixal clitics whose attachment is determined by an independent alignment constraint. From this proposal it follows that clitic clusters can be nicely derived as composed units and positioned into preverbal or postverbal position. Crucially, we capture that a) the order of clitics inside the cluster is not dependent on the direction of attachment of the cluster and b) that the unit appearing before or after the host is generated by the same set of realisation rules.

d) Alignment constraints specify the directional and hierarchical properties associated with the composed affixal unit in (2). In effect, one of the roles of morphology is to differentiate between the pure suffix status of enclitics and the phrasal affix status of proclitics. Two placement rules have been proposed for EP, which attach affixal clitics to the left of a Vº phrasal node (for proclisis, 3a) and to the right of a verbal stem (for enclisis, 3b).

(3) a. < <aff,aff >, Vº >, if Vf [Restricted: yes]
b. < stem, <aff,aff > >, elsewhere

This style of analysis can be used to account for other clitic systems, such as clitic pronouns in other Romance languages which attach uniformly in the morphology, or typical Wackernagel clitics as in Bosnian/Croatian/Serbian which attach exclusively as phrasal affixes. In fact, the Paradigm Function model can successfully capture instances in which clitics cluster constitute affixal strings attached in

the domain of a morphological word, phrasal affix strings attached in the domain of a syntactic construction, or both.

The hypothesis that exponence should be separated from placement was motivated essentially with data on cliticisation which raised the question of how to position the same affix in preverbal and postverbal position. As alluded to before, an insight by Stump (1993) on ambifixal exponents provided the conceptual basis upon which to consider the 'separation' approach. The idea of deriving sequences of affixes through function composition, suggested in Spencer (2000), made it possible to treat sequences of affixes as composed affixal sequences prior to their placement.

During the period within which this thesis was elaborated, various papers appeared which extend this approach to other instances of affixation, including ordinary inflection and derivation (Luís&Spencer 2004, Spencer ms, Otoguro 2003, Spencer 2003). A detailed exposition of the arguments supporting a revised model of Paradigm-Function Morphology for inflection and derivation are presented in Spencer (ms).

The morphology-syntax interaction

Formally, the inflectional approach to clitic pronouns has been embedded into Lexical-Functional Syntax (Bresnan 2001). In other words, a combined framework is adopted which incorporates the realisational model of PFM into the architecture of Lexical-Functional Grammar (Kaplan&Bresnan 1982), building on insights by Sadler&Spencer (2001), Luís&Sadler (2003), Sadler&Nordlinger (2003).

Among the implication of placing realisational morphology in correspondence with LFG-syntax is the fact that affixes can no longer be viewed as sub-lexical entries. So, whereas in classical LFG f-structure are associated directly with affixes, in 'realisational' LFG an explicit interface between morphological structures and the syntax must be assumed which explores the correspondence between f-structure information and morphological features (Kaplan&Butt 2002, among other). In the present thesis, a brief sketch of how such a correspondence might look was

illustrated with a set of mapping rules which associate the f-strtucture information associated with pronominal clitics to the morphological features (for cases illustrating more elaborate morphology-syntax correspondences, cf. Sadler&Spencer 2001, Luís&Spencer in press, Sadler&Nordlinger 2003, among other).

Within this combined PFM/LFG framework, it has been shown that the syntactic nature of the factors conditioning clitic placement does not invalidate an inflectional analysis of cliticisation (contrary to previous claims). The present analysis has assumed that if all proclitic constructions are mapped onto an abstract functional feature and if that f-feature is placed in correspondence with the morphological features triggering placement in the morphology, the fact that proclisis is dependent on specific elements in the clause is captured naturally without unnecessary extensions to the theory. A more detailed analysis of the mapping between f-structure and morphology remains to be explored. As future research, I shall examine the c-structural positions of each one of the proclitic triggers and explore within LFG how precedence relations between the trigger and the verb can be encoded into f-structure information and made directly available to the morphology[86].

[86] An account along these lines is currently being investigated by Luís&Otoguro (in prep) within Lexical-Functional Grammar.

Appendix A

Fragment of verbal morphology of EP

1. Some verbal paradigms

(the sign '=' indicates stem-internal boundaries and the sign '-' separates tense and agreement suffixes).

(1)Imperfect paradigm

tense	agr	stem	affixal positions		
			I	II	
Imperfect	1SG	lev=a	-va		'I carried'
	2SG	lev=a	-va	-s	'you carried'
	3SG	lev=a	-va		's/he carried'
	1PL	lev=a	-va	-mos	'we carried'
	2PL	lev=a	-ve	-is	'you.pl carried'
	3PL	lev=a	-va	-/N/	'they carried'

(2) Future paradigm

tense	agr	stem	affixal positions		
			I	II	
Future	1SG	lev=a=r	-e	-i	'I will carry'
	2SG	lev=a=r	-á	-s	'you will carry'
	3SG	lev=a=r	-á		's/he will carry'
	1PL	lev=a=r	-e	-mos	'we will carry'
	2PL	lev=a=r	-e	-is	'you.pl will carry'
	3PL	lev=a=r	-a	-/N/	'they will carry'

(3)Conditional paradigm

tense	agr	stem	affixal positions		
			I	II	
Conditional	1SG	lev=a=r	-ia		'I would carry'
	2SG	lev=a=r	-ia	-s	'you would carry'
	3SG	lev=a=r	-ia		's/he would carry'
	1PL	lev=a=r	-ia	-mos	'we would carry'
	2PL	lev=a=r	-ia	-is	'you.pl would carry'
	3PL	lev=a=r	-ia	-/N/	'they would carry'

2- Rule Blocks

Based on the paradigms above, I propose the two following rule blocks. Block I contains realisation rules which introduce tense (and agreement) suffixes. Block II contains rules which realise exponents of agreement.

(4)

Block I

RR Ia, {TENSE:Impf}, V (<X, σ>) = def <X-*va*,σ>
RR Ib, {TENSE:Impf, PERSON: 2, NUMBER: Pl}, V (<X, σ> = def <X-*ve*,σ>
RR Ic, {TENSE:Future}, V (<X, σ>) = def <X-*e*,σ>
RR Id, {TENSE:Future PERSON: 2/3, NUMBER: SG }, V (<X, σ>) = def <X-*a*,σ>
RR Ife {TENSE:Cond}, V (<X, σ>) = def <X-*ia*,σ>

Block II

RR IIa, { PERSON: 2, NUMBER: SG }, V (<X, σ>) = def <X-*s*,σ>
RR IIb, { PERSON: 1, NUMBER: Pl}, V (<X, σ>) = def <X-*mos*,σ>
RR IIc, { PERSON: 2, NUMBER: Pl}, V (<X, σ>) = def <X-*is*,σ>
RR IId, { PERSON: 3, NUMBER: Pl}, V (<X, σ>) = def <X-*N*,σ>
RR IIe, {TENSE:Future, PERSON: 1, NUMBER: SG}, V (<X, σ>) = def <X-*i*,σ>

3- Indexed stems

The realisation rules given in (4) take as input one of the following indexed stems:

(5)

a. $Stem_1$ = Root + *-a*, *-e*, *-I* (e.g., *lev*+*a*, *beb*+*e*, *ment*+*i*).
b. $Stem_2$ = $Stem_1$ + *-r* (e.g., *leva*+*r*), except for irregular forms.

Stems might be derived through specific realisation rules, namely morphomic rules (using insights from Aronoff 1994). These rules are different from ordinary RRs because they do not realise morphosyntactic features but instead derive 'meaningless' forms (cf. Chapter 6, section 6.2). Let us therefore assuming that the 'morphomic' level of morphology (Aronoff 1994) provides morphemic rules; let us also assume that morphemic rules (MR) are organised into ordered

morphemic rule blocks. Given this small set of assumptions, this is how the stems given in (5) might be derived within PFM:

Rule Block I provides MR for the derivation of first, second and third class conjugation stems (5a). These MR rules take as input the lexemic root (e.g., *lev*- of LEVAR 'take', or *beb*- of BEBER 'drink' or *men*t- of MENTIR 'lie') and, depending on the conjugation class of the lexeme, they associate the root with one of the three theme vowels (i.e., *-a*, *-e*, or -i). This means that a) $MR_{I\text{-}1}$ from Rule Block I realises the theme vowel *-a* for lexemes of the first conjugation, b) $MR_{I\text{-}2}$ from Rule Block I realises the theme vowel *-e* for lexemes of the second conjugation, and c) $MR_{I\text{-}3}$ from Rule Block I realises the theme vowel *-i* for lexemes of the third conjugation.

Likewise, we would need another morphemic Rule Block, namely Rule Block II, for the derivation of the infinitival stem in (5b), which would apply after Rule Block I. Evidently, these two stems do not exhaust the range of stems that are necessary for the whole verbal paradigm of EP. We leave a more detailed and formalised treatment of this topic for future research (cf. Spencer (ms), Popova (to appear), on stem formation within PFM).

Appendix B

The clitic inventory

In formulating the RRs, it is important to observe that the EP clitic system displays various instances of syncretism. As the table shows, Gender is only relevant for 3rd accusative clitics; Number fails to combine with 3rd Reflexives and, none of the 1st and 2nd person clitics is specified for Case. Case syncretism then is found with all 1st person and all 2nd person clitics, both in the singular and plural forms. Also, number syncretism is found on 3rd person reflexives, given that there is just one form for both the singular and plural clitics.

(1) EP clitic systems

<table>
<tr><th></th><th>REFLEXIVE</th><th>DATIVE</th><th>ACCUSATIVE</th></tr>
<tr><td>1.Sg.</td><td>me</td><td>me</td><td>me</td></tr>
<tr><td>2.Sg.</td><td>te</td><td>te</td><td>te</td></tr>
<tr><td>3.Sg.Masc.</td><td rowspan="2">se</td><td rowspan="2">lhe</td><td>o</td></tr>
<tr><td>3.Sg.Fem.</td><td>a</td></tr>
<tr><td>1.Pl.</td><td>nos</td><td>nos</td><td>nos</td></tr>
<tr><td>2.Pl.</td><td>vos</td><td>vos</td><td>vos</td></tr>
<tr><td>3.Pl.Masc</td><td rowspan="2">se</td><td rowspan="2">lhes</td><td>os</td></tr>
<tr><td>3.Pl.Fem.</td><td>as</td></tr>
</table>

One insightful way of capturing this syncretism is to assume that there is only one RRs for each person and number combination, thus leaving Case unspecified[87]. The rules in (2), then, distinguish between the values for Person and Number, but not Case. This means that each rule can realise any value for Case and Reflexivity.

[87] The format for RRs adopted here has been introduced in the revised analysis in Chapter 9. This proposal assumes a separation between the realization and the placement of affixes (Spencer ms, Luís&Spencer in press, Luís 2003a). The phonological form of the affixal clitics is discussed in chapter 7.

(2) RRs for 1st person and 2nd person clitics (provisional)

a. $RR_{B1, \{ \text{PERSON: 1, NUMBER: Sg} \}, V}(\sigma)$ $=_{def}$ [*mE*, σ]
b. $RR_{B2, \{ \text{PERSON: 2, NUMBER: Sg} \}, V}(\sigma)$ $=_{def}$ [*tE*, σ]
c. $RR_{B3, \{ \text{PERSON: 1, NUMBER: Pl} \}, V}(\sigma)$ $=_{def}$ [*nos*, σ]
d. $RR_{B4, \{ \text{PERSON: 2, NUMBER: Pl} \}, V}(\sigma)$ $=_{def}$ [*vos*, σ]

Whereas in Macedonian clitics attach to an inflected verb form, and therefore RRs are defined over inflected verbs, in EP RRs apply over the stem of a lexeme. I therefore adopt the standard format, where X stands for an underived base. So, the left hand side of the rule specifies a) the features it realises and b) the morphological base X (i.e. root or stem) to which the RR applies. The right hand side of the rule specifies the result of applying the rule to X.

The inventory in (1) also shows that 3rd person Reflexive clitics are undetermined for Number. To capture that *se* can be the exponent of both singular and plural forms, I propose the rule in (3) (cf. Miller&Sag 1997, for similar view):

(3) Feature set for 3rd reflexive clitics (provisional)

$RR_{A1, \{\text{PERSON:3, REFL:+}\}, V}(\sigma)$ $=_{def}$ [*se*, σ]

Moving to the representation of 3rd person clitics, it is clear that a quite different set of feature combinations must be provided to capture the paradigmatic contrasts. In this case, Case and Reflexivity play a distinctive role, given that 3rd reflexive, dative and accusative clitics are all distinctly differentiated. Gender also has a role to play in the distinction between singular and plural accusative clitics, unlike for dative or reflexive clitics. As seen above, reflexive clitics don't differentiate Number features either:

(4) Inventory for 3rd person clitics

	REFLEXIVE	DATIVE	ACCUSATIVE
3.Sg.Masc.			o
3.Sg.Fem.		lhe	as
3.Pl.Masc	se		os

3.Pl.Fem.		lhes	as

The RRs rules proposed for 3rd person clitics are summarised in (5). Dative clitics encode Number distinctions and must be realised by two separate RRs: one for 3rd person singular (5a) and one for 3rd person plural (5b). Specifying these RRs for Number indicates that 3rd singular and 3rd plural datives correspond to distinct affixes. Finally, accusative clitics offer a richer inventory, since they also differentiate between masculine and feminine forms. RRs for accusatives thus must be enriched with values for Gender, in addition to the values for Case, Person and Number. The representation of the set of accusative clitics is given in (5c-f).

(5) Feature sets for 3rd person clitics (provisional)

a. $RR_{B5, \{CASE: Dat, PERSON:3, NUMBER: Sg\}, V}(\sigma)$ $=_{def}$ [lhe, σ]
b. $RR_{B6, \{CASE: Dat, PERSON:3, NUMBER: Pl\}, V}(\sigma)$ $=_{def}$ [lhes, σ]
c. $RR_{C1, \{CASE: Acc, PERSON:3, NUMBER: Sg, GENDER: Masc\}, V}(\sigma)$ $=_{def}$ [LNo, σ]
d. $RR_{C2, \{CASE: Acc, PERSON:3, NUMBER: Sg, GENDER: Fem\}, V}(\sigma)$ $=_{def}$ [LNa, σ]
e. $RR_{C3, \{CASE: Acc, PERSON:3, NUMBER: Pl, GENDER: Masc\}, V}(\sigma)$ $=_{def}$ [LNos, σ]
f. $RR_{C4, \{CASE: Acc, PERSON:3, NUMBER: Pl, GENDER: Fem\}, V}(\sigma)$ $=_{def}$ [LNas σ]

The complete inventory of clitic features that are expressed by EP pronominal clitics and the permissible values for these features are summarised below:

(6) Clitic features

CASE: Accusative, Dative
REFLEXIVITY: yes, no
PERSON: 1,2,3
NUMBER: Singular, Plural
GENDER: Masculine, Feminine

The features include Case, Reflexivity, Person, Number, and also Gender. The permissible values for these features are a) 'accusative' and 'dative' for the feature Case, b) 'yes' and 'no' for the feature Reflexivity, c) '1st', '2nd' and '3rd' for Person, d) 'singular' and 'plural' for Number, e) 'masculine' and 'feminine' for Gender.

References

Ackerman, Farrell and Philippe Lesourd (1997) Towards a Lexical representation of phrasal predicates. In: A. Alsina et al. (eds.) *Complex Predicate*. Stanford: CSLI.

Ackerman, Farrell and Gregory T. Stump (to appear) Paradigms and periphrastic expression: a study in realization-based lexicalism. To appear in Louisa Sadler and Andrew Spencer (eds.) *Projecting Morphology*. Stanford University: Center for the Study of Language and Information.

Ackerman, Farrell and G. Webelhuth (1998) *A Theory of Complex Predicates*. Stanford: CSLI.

Aguado, Miguel and Christian Lehmann (1989) Zur Grammatikalisierung der Klitika im Katalanischen. In: Wolfgang Raible (ed.) *Romanistik Sprachtypologie und Universalienforschung. Beiträge zum Freiburger Romanistentag 1987*. Tübingen: Gunter Narr, pp. 151-161.

Ambar, Manuela (1992) *Para uma Síntaxe em Português da Inversão Sujeito-Verbo.* Lisboa: Ed. Colibri.

Anderson, Stephen R. (1992) *A-Morphous Morphology*. Cambridge Studies in Linguistics 62. Cambridge: CUP.

Anderson, Stephen R. (1993) Wackernagel's Revenge: Clitics, Morphology and the Syntax of second Position. *Language* 69, pp. 68-96.

Anderson, Stephen R. (1995) Rules and Constraints in Describing the Morphology of Phrases. *CLS Parasession on Clitics*. Chicago Linguistic Society, pp. 15-31.

Anderson, Stephen R. (1996). How to Put Your Clitics in Their Place. *The Linguistic Review* 13, pp. 165-191.

Anderson, Stephen R. (2000). Towards an Optimal Account of Second Position Phenomena. In: J. Dekkers, et al. [eds.], *Optimality Theory: Syntax, Phonology and Acquisition* (Oxford University Press). pp. 302-333. Also in Lizanne Kaiser, ed., *Studies in the Morphosyntax of Clitics* (Y.A.L.E. working papers, vol. 1, 1997)

Andrews, A. (1990) Unification and Morphological Blocking. *Natural Language & Linguistic Theory* 8, pp. 507-557.

Aronoff, M.(1976) *Word Formation in Generative Grammar*. MIT Press, Cambridge MA.

Aronoff, Mark (1995) *Morphology by Itself*. Cambridge, Mass.: The MIT Press.

Aronoff, Mark and S. N. Sridhar (1983) Morphological Levels in English and Kannada or Atarizing Reagan. In: J. F. Richardson et. al. (eds.) *Papers from the Parasession on The Interfaces of Phonology, Morphology and Syntax*. Chicago, CLS.

Auger, Julie (1994) *Pronominal Clitics in Québec Colloquial French: A Morphological Analysis*. Phd Diss. Philadelphia: University of Philadelphia.

Barbosa, Pilar (1996) Clitic Placement in European Portuguese and the Position of Subjects. In: A. Halpern and A. Zwicky (eds.) *Approaching Second: Second Position Clitics and Related Phenomena*. Stanford: CSLI.

Bechara, Evanildo 2000 *Moderna Gramática Portuguesa*. Rio de Janeiro: YHL.

Belleti, Adriana (1999) Italian/Romance Clitics: Structure and Derivation. In: Henk van Riemsdijk, ed., *Language typology: Clitics in the European Languages*. Berlin: Mouton de Gruyter, pp. 543-579.

Berendsen, Egon (1986) *The Phonology of Cliticization*. Dordrecht: Foris.

Bochner, Harry (1993) *Simplicity in Generative Morphology.* Publications in Language Series 37. Berlin. New York: Mouton de Gruyter.

Bonet, Eulàlia (1991) *Morphology After Syntax: Pronominal Clitics in Romance*. PhD Thesis: MIT.

Booij, Geert (1983) Principles and Parameters in Prosodic Phonology. In: *Linguistics* 21, pp. 249-280.

Booij, Geert (1995) *The Phonology of Dutch*. Oxford: Clarendon Press.

Booij, Geert (1996) Cliticization as Prosodic Integration: The Case of Dutch. *The Linguistic Review* 13, pp. 219-242.

Börjars, K., N., Vincent and C. Chapman (1996) Paradigms, Periphrasis and Pronominal Inflection: a feature-based account. In: G. Booij and J. van Marle, eds., *Yearbook of Morphology 1996.* Kluwer: Dordrecht, 155-180.

Börjars, Kersti (2003) Functional categories and position in the noun phrase. Paper presented at the *York-Essex Morphology Meeting*, University of York.

Bresnan, Joan (2000) Optimal Syntax. In: Joost Dekkers, Frank van der Leeuw and Jeroen van der Weijer (eds.) *Optimality Theory: Syntax, Phonology, and Acquisition*. Oxford: OUP.

Bresnan, Joan (2001) *Lexical Functional Syntax*. Oxford: Blackwell.

Bresnan, Joan and Sam Mchombo (1987) Topic, Pronoun and Agreement in Chichewa. *Language* 63.4, pp. 741-782.

Brines, Natalia (2001) *Spanish Clitics and Argument Reduction Processes*. PhD Thesis, Essex University.

Cardinaletti, Anna and Michal Stark (1999) The Typology of Structural Deficiency: On the Three Grammatical Clases. In: *Clitics in the Languages of Europe*. Berlin/New York: Mouton de Gruyter, pp. 145-233.

Carstairs-McCarthy, Andrew (1987) *Allomorphy in Inflexion.* London: Croom Helm.

Costa, João (2000) Word Order and Discourse-Configurationality in Europena Portuguese. In: João Costa (ed.) *Portuguese Syntax. New Comparative Studies*. Oxford: Oxford University Press, pp. 95-115.

Chomsky, Noam (1970) Remarks on nominalisation. In R. Jacobs & P. Rosenbaum, eds. *Readings in Transformational Gramnmar*. Waltham: Ginn, pp. 184-221.

Chomsky, N. (1981). *Lectures on government and binding*, Foris Publications, Dordrecht.

Chomsky, Noam, & Howard Lasnik (1993) Principles and parameters theory. In: Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld and Theo Vennemann (eds.) *Handbook Syntax*. Berlin: Walter de Gruyter.

Chomsky, N. (1995) *The Minimalist Program.* Cambridge, MA: MIT Press.

Crysmann, Berthold (1997) Cliticization in European Portuguese using Parallel Morpho-syntactic Constraints. In: Miriam Butt and Tracy H. King (eds.) *Proceedings of the LFG97 Conference*. Stanford, CA: Center for the Study of Language and Information.

Crysmann, Berthold (2000) Syntactic Transparency of Pronominal Affixes. In: Ronnie Cann, Claire Grover, and Philip Miller (eds.) *Grammatical Interfaces in HPSG, Studies in Constraint-Based Lexicalism.* CSLI Publications, Stanford, pp. 77-96.

Crysmann, Berthold (1999) Licensing Proclisis in European Portuguese. In: Francis Corblin, Carmen Dobrovie-Sorin and Jean Marie Marandin (eds.) *Empirical Issues in Formal Syntax and Semantics 2. Selected Proceedings form the Colloque de Syntaxe et Sémantique de Paris*. Thesus, The Hague, pp. 255-276

Crysmann, Berthold (2002) *Constraint-based Coanalysis*. PhD Thesis, Universität des Saarlandes & DFKI Gmbh.

Cummins, S. and Y. Roberge (1994) A Morphosyntactic Analysis of Romance Clitic Constructions. In Michael Mazzola (ed.) *Issues and Theory in Romance Linguistics.* Georgetown University Press: Washington. pp. 237-255. 1994.

Cunha, Celso and Lindley Cintra (1987) *Nova Gramática do Português Contemporâneo*. Lisboa: Edições Sá da Costa.

Dalrymple, Mary (2001) *Lexical Functional Grammar*. Syntax and Semantics Series. Volume 34. Academic Press.

Duarte, Inês et al. (1995) Specificity of European Portuguese Clitics in Romance. In: Hub Faria et al. (eds.) *Studies on the Acquisition of Portuguese*. Lisboa: APL/Colibri.

Duarte, Inês and Gabriela Matos (2000) Romance Clitics and the Minimalist Program. In: João Costa (ed.) *Portuguese Syntax. New Comparative Studies*. Oxford: Oxford University Press, pp. 116-142.

Emonds, J. (1975) A transformational analysis of French clitics without positive output constraints. In: *Linguistic Analysis* 1.1, pp. 3-24.

Falk, Yehuda (2001) *Lexical-Functional Grammar*. Stanford, CA: CSLI Publications.

Fernández Soriano, O. (1993): *Los pronombres átonos*. Madrid: Taurus Universitaria

Frota, Sónia (1996) Prosodic Phrases and European Portuguese: in Search of Evidence, In: A. Bisetti et al. (eds.) *Proceedings of Console III*. Leiden: Sole.

Gerlach, Birgit (1998) Optimale Klitiksequenzen. *Zeitschift für Sprachwissenschaft* 17, pp. 35-91.

Gerlach, Birgit (2001a) *Clitics Between Syntax and Lexicon*. PhD thesis, Heinrich-Heine Universität, Düsseldorf.

Gerlach, Birgit (2001b) Neutralisation in European Portuguese. Paper presented at the Heinrich-Heine Universität.

Grimshaw, Jane (1982) Romance Reflexive Clitics. In: Joan Bresnam (ed.) *The Mental Representation of Grammatical Relations*. MIT Press, Cambridge, MA, pp. 173-281.

Grimshaw, Jane (1997) The Best Clitic: Constraint Conflict in Morphosyntax. In: Liliane Haegemann (ed.) *Elements of Grammar*. Dordrecht: Kluwer, pp. 169-196.

Haegeman, Liliane (1992) *Generative Syntax: theory and description. A case study from West Flemish*. Cambridge: CUP.

Halpern, Aaron (1995) *On the Placement and Morphology of Clitics*. Stanford: CSLI publications.

Halpern, Aaron (1998) Clitics. In: Andrew Spencer and Arnold M. Zwicky (eds.) *The Handbook of Morphology*. Oxford: Blackwell.

Hammond, Michael (1999) *The Phonology of English: A Prosodic Optimality-Theoretic Approach*. Oxford. OUP.

Harris, Alice C. (2000) Where in the world is the Udi Clitic? *Language* 73 (3), pp. 593-616.

Harris, Alice C. (2002) *Endoclitics and the Origins of Udi Morphosyntax*. Oxford: Oxford University Press.

Harris, James (1994) Projection and Edge Marking in the Computation of Stress in Spanish. In John Goldsmith (ed) *A Handbook of Phonological Theory*. 867-887, Oxford: Basil Blackwell.

Harris, Nigel and Nigel Vincent (1988) *The Romance Languages*. London: Croom Helm.

Hayes, B. (1990) Precompiled Phrasal Phonology. In: S. Inkelas and D. Zec (eds) *The Phonology-Syntax Connection*. The University of Chicago Press, Chicago, pp. 127-143.

Hauge, Kjetil (1976) *The Word Order of Predicate Clitics in Bulgarian*. Meddelelser, Slavisk-Baltisk Institutt, Vol. 10. Oslo, Universitetet i Oslo.

Heap, David (2000) Morphological Complexity and Spanish Object Clitic Variation. *30th Linguistic Symposium on Romance Languages*. (To appear in C. Wiltshire & J. Camps (eds), Romance Phonology and Variation. Selected papers from the Amsterdam: John Benjamins)

Heine, Bernd, Claudi, Ulrike, Huennemeyer, Friederike (1991) *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.

Hopper, Paul J., Traugott, Elizabeth Closs (1993) *Grammaticalization*. Cambridge: Cambridge University Press.

Joseph, B. D. and Smirniotopoulos, J. C. (1993) The morphosyntax of the Modern Greek verb as morphology and not syntax. *Linguistic Inquiry*, 24(2):388-398.

Kanerva, Jonni (1987) Morphological Integrity and Syntax: The Evidence from Finnish Possessive Suffixes. *Language* 63, pp. 498-521.

Kaplan, Ron and Joan Bresnan (1982) Lexical-functional Grammar: a formal system for grammatical representation. In: Joan Bresnam (ed.) *The Mental Representation of Grammatical Relations*. MIT Press, Cambridge, MA, pp. 173-281.

Kaplan, Ron and Miriam Butt (2002) The Morphology-Syntax Interface in LFG. Unpublished Handout, *LFG02*.

Kathol, Andreas (2000) *Linear Syntax*. Oxford: Oxford University Press.

Kayne, Richard (1975) *French Syntax: The Transfomational Cycle*. Cambridge MA: MIT Press.

Kayne, Richard (1991) Romance Clitics, Verb Movement, and PRO. *Linguistic Inquiry* 22 (4), pp. 647-686.

Klavans, Judith (1980) *Some Problems in a Theory of Clitics*. Bloomington: Indiana University Linguistics Club.

Klavans, Judith (1985) The Independence of Syntax and Phonology in Cliticisation. *Language*, 61, pp. 95-120.

Lapointe, Steven (1980) *A Theory of Grammatical Agreement*. PhD. UMASS Amherst.

Lecarme, J. (1991) Focus en Somali: Syntaxe et interpretation. *Linguistique Africaine* 7, pp. 275-309.

Leeuw, Frank van der (1997) Clitics: Prosodic Studies. The Hague: Holland Academic Graphics.

Legendre, Géraldine (1996) *Clitics, Verb (Non-) Movement, and Optimality in Bulgarian*. Technical Report JHU-CogSci-96.5. Baltimore: John Benjamins Publishing Company.

Legendre, Géraldine (1998) Second Position Clitics in a Verb-Second Language: Conflict Resolution in Macedonian. In: S. Avrutin and D. Jonas (eds.) *Proceedings of ESCOL 1997*. Cornell University: GLC Publications, pp. 139-149.

Legendre, Géraldine (2000a) Positioning Romanian Verbal Clitics at PF: an Optimality-Theoretic Analysis. In: Birgit Gerlach and Janet Grijzenhout (eds.) *Clitics in Phonology, Morphology, and Syntax*. John Benjamins Publishing Company: Amsterdam/Philadelphia, pp. 219-254.

Legendre, Géraldine (2000b) Morphological and Prosodic Alignment of Bulgarian Clitics. In: Joost Dekkers, Frank van der Leuuw and Jerob van der Weijer (eds.) *Optimality Theory: Syntax, Phonology, and Acquisition*. Oxford: OUP, pp. 423-462.

Luís, Ana (1997) English reduced auxilaries. In: Markve, S., E. Pennings and A. Tsukiashi (eds.) *Proceedings of the Sixth Manchester Postgraduate Linguistics Conference*. Department of Linguistics, University of Manchester, pp. 213-238.

Luís, Ana (2000) The Grammaticalisation of Object Pronouns. *Akten des 35. Linguistisches Kolloquiums, University of Innsbruck*. Volume 2. Peter Lang: Frankfurt, pp. 85-95.

Luís, Ana (2001a) The Inflectional Derivation of Clusters. Presented at the *Autumn Meeting of the Linguistic Association of Great Britain*, University of Reading.

Luís, Ana (2001b) Function Words, Prosodic Structure and Morphological Dependency. In: Marjo van Koppen, Joanna Sio, Mark de Vos (eds.) *Proceedings of Console X, 2001*. Leiden: SOLE, pp. 157-171.

Luís, Ana (2003a) On the Inflectional Status of Pronominal Clitics. M. Cazzoli-Goeta, M. Heather and S. Pourcel (eds.) *Durham Working Papers in Linguistics*, volume 9, pp. 25-38.

Luís, Ana (2003b) The Morphophonology of Pronominal Affixes in European Portuguese. In: *Proceedings of the Western Conference of Linguistics 2002*. The University of British Columbia: Vancouver.

Luís, Ana (to appear). Cluster Formation and Placement Asymmetries. In: *Proceedings of the 38th Regional Meeting of the Chicago Linguistic Society, 2002*. Chicago: Chicago University Press.

Luís, Ana and Otoguro Otoguro (in prep.) The contexts triggering proclisis in European Portuguese and their effect on clitic placement. University of Coimbra/University of Essex.

Luís, Ana and Andrew Spencer (in press) A Paradigm Function Account of 'mesoclisis' in European Portuguese. In: G. Booij and J. van Marle, eds., *Yearbook of Morphology.* Kluwer: Dordrecht.

Luís, Ana and Louisa Sadler (2003) Object Clitic and marked Morphology. In: Beyssade, C., O. Bonami, P. Hofherr and F. Corblin (eds.) *Empirical Isues in Formal Syntax and Semantics 4*. Paris: Presses de l'Université de Paris-Sorbonne, pp. 133-153.

Luutonen, Jorma (1997) *The variation of morpheme order in Mari declension.* Helsinki: Suomalais-ugrilainen seura.

Madeira, Ana Maria (1991) On clitic placement in European Portuguese. *University College London WPL* 4, pp. 95-122.

Madeira, Ana Maria (1993) Clitic-Second in European Portuguese. *Probus* 5, pp. 154-175.

Martins, Ana Maria (1994) *Clíticos na História do Português*. PhD Thesis. University of Lisbon.

Mateus, M. H. M. (1975) *Aspectos da Fonologia Portuguesa*. Lisboa: INIC.

Mateus, M. H. M. (1983) O acento da palavra em português: uma nova proposta. *Boletim de Filologia* 28, pp. 95-120.

Mateus, M. H. M., A. Brito, I, Duarte & I. Faria (1989) *Gramática da Língua Portuguesa*. Lisboa : Ed. Caminho.

Mateus, M. H. & E. d'Andrade (2000) *The Phonology of Portuguese*. Oxford: OUP.

Matos, Gabriela (1997) Configurações sintácticas em estruturas de colocação simultânea de clítico. In : Brito, A. et al. (eds*.) Sentido que a vida faz. Estudos para Óscar Lopes*. Porto: Campo das Letras, pp. 705-717.

Matos, Gabriela (2000) Across-the-Board clitic placement in Romance languages. *Probus* 12-2, Page 229-260

Matthews, Peter (1972) *Inflectional Morphology*. Cambridge: CUP.

Matthews, P. H. 1991.*Morphology.* 2nd edition. Cambridge, England: Cambridge University.

McConvell, P. (1996) The Functions of Split-Wackernagel Clitic Systems: Pronominal Clitics in the Ngumpin Languages (Pama-Nyungan family, Northern Australia). In: A. Halpern and A. Zwicky (eds.) Approaching Second Position Clitics and Related Phenomena. Stanford: CSLI.

Menendez Pidal, R. (1904) *Manual de gramática histórica española*. Madrid: Espasa-Calpe.

Miller, Philip (1992) *Clitics and Constituents in Phrase Structure Grammar.* Garland Publishers.

Miller, P. and I. Sag (1997) French Clitic Movement without Clitics or Movement. *Natural Language and Linguistic Theory* 15 (3), pp. 573-639.

Monachesi, Paola (1994) Towards a typology of Italian clitics. In: K. Beals, J. Denton, R. Knippen, L. Melnar, H. Susuki and E. Zeinfeld (eds.) *Proceedings of the 30th Regional Meeting of the Chicago Linguistic Society*. Chicago, pp. 266-280.

Monachesi, Paola (1998) The morphosyntax of Romanian Cliticization. In: P. Coppen, H. van Halteren and L. Teunissen (eds.) *Proceedings of Computacional Linguistics in The Netherlands 1997*. Rodopi, Amstardam, pp. 99-118.

Monachesi, Paola (1999) *A Lexical Approach to Italian Cliticization*. CSLI, Stanford.

Monachesi, Paola (2000) Linearization properties of Romanian Clitics. In: Birgit Gerlach and Janet Grijzenhout (eds.) *Clitics in Phonology, Morphology, and Syntax*. Amstardam/Philadelphia: John Benjamins.

Nespor, Marina and Irene Vogel (1986) *Prosodic Phonology*. Dordrecht: Foris.

Nevis, Joel (1985) *Finnish Particle Clitics and General Clitic Theory*. PhD, The Ohio State University.

Nishida, Chiyo (1987) Interplay between Syntax and Morphology: A Lexical Analysis of Inflection and Cliticization in Spanish. *ESCOL '87*, The Ohio State University.

Nordlinger, Rachel (1998) *Constructive Case: Evidence from Australian languages*. Stanford: CSLI Publications

Nordlinger, Rachel and Joan Bresnan. To appear. Lexical-Functional Grammar: interactions between morphology and syntax. In R. Borsley and K. Börjars (eds) *Non-transformational syntax: a guide to current models*. Oxford: Blackwells.

Noyer, Rolf (1994) Mobile affixes in Huave: Optimality and Morphological Wellformdess. In: Erin Duncan, Donka Farkas and Philip Saelti (eds.) *Proceedings of the Twelph West Coast Conference on Formal Linguistics*. Santa Cruz: University of California, pp. 67-82.

O'Connor, R. (2002) Clitics as Phrasal Affixation in Constructive Morphology. In: Miriam Butt and T. King (eds.) *Proceedings of the LFG02*. Stanford: CSLI Publications. (Availabe http: csli-publications.stanford.edu/)

Orgun, O. (1996) *Sign-Based Morphology and Phonology*. Unpublished PhD Thesis, University of California, Berkeley.

Orgun, O. and S. Inkelas (1998) Level (Non)ordering in Recursive Morphology: Evidence from Turkish. In S. Lapointe (ed.), *Morphology and its Relations to Syntax and Phonology.* Stanford: CSLI Publications. 360-392.

Otoguro, Otoguro (2003) Focus Clitics and Discourse Information Spreading. In Miriam Butt and Tracy Holloway King (eds.) *Proceedings of the LFG03 Conference*, University at Albany, State University of New York, July 2003. CSLI Publications, http://csli-publications.stanford.edu/

Peperkamp, Sharon (1997) *Prosodic Words*. The Hague: Holland Academic Graphics.

Perlmutter, David (1970) Surface Structure Constraints in Syntax. *Linguistic Inquiry* 1, pp. 187-255.

Pereira, Isabel (1999) *O acento de palavra em português. Uma análise métrica*. Coimbra: Universidade de Coimbra.

Pollard, C. and I. Sag (1994) *Head-Driven Phrase Structure Grammar*. Stanford CA: CSLI and University of Chicago Press.

Popescu, Alexandra (2000) The Morphophonology of the Romanian Clitic Sequence. *Lingua* 110, pp. 773-799.

Popova, Gergana (to appear) Bulgarian Aspect in Paradigm-Function Morphology. In: A. Galani and R. Otoguro (eds.) *Proceedings of the First York-Essex Morphology Meeting*. University of York.

Poser, William (1985) *Cliticization to NP and Lexical Phonology*. In: J. Goldberg et. al. (eds.) Proceedings of WCCFL 4. Stanford: CSLI Publications.

Prince, Alan and Paul Smolensky (1993) *Optimality Theory: Constraint Interaction in Generative Grammar*. Cambridge Mass.: MIT Press.

Raposo, Eduardo (2000) Clitic Positions and Verb Movement. In: João Costa (ed.) *Portuguese Syntax. New Comparative Studies*. Oxford: Oxford University Press, pp. 266-297.

Riemsdijk, Henk van (1999) Clitics: A State Of The Art Report. In: Henk van Riemsdijk, ed., *Language typology: Clitics in the European Languages*. Berlin: Mouton de Gruyter.

Roberts, Ian and Anna Cardinaletti (1991) Clause Structure and X-Second. (Available at http://kiri.ling.cam.ac.uk/roberts/xsecond.rtf)

Sadler, Louisa (1997a) Clitics and the Structure-Function Mapping. In: Miriam Butt and T. King (eds.) *Proceedings of the LFG97*. Stanford: CSLI Publications.

Sadler, Louisa (1997b) English Auxiliaries as Tense Inflections. Special Issue of Essex Research Reports

Sadler, L. and D. Arnold (1994) Prenominal Adjectives and the Phrasal/lexical Distinction. *Journal of Linguistics* 30, pp. 187-226.

Sadler, Louisa and Rachel Nordlinger (2002) Case Stacking in Realisational Morphology. Paper presented as LFG02. (Available at http://privateweb.essex.ac./~louisa)

Sadler, Louisa and Rachel Nordlinger (2003) Relating Morphology to Syntax. (Available at http://privateweb.essex.ac./~louisa) (To appear in Louisa Sadler and Andrew Spencer (eds.) *Projecting Morphology*. Stanford: CSLI Publications.)

Sadler, Louisa and Andrew Spencer (2001) Syntax as an Exponent of Morphological features. In: Geert Booij and Jaap van Marle (eds.) *Yearbook of Morphology 2000*, Dordrecht: Kluwer Academic Publishers, pp. 71-96.

Sadock, Jerrold (1991) *Autolexical Syntax: A Theory of Parallel Grammatical Representations*. Chicago, University of Chicago Press.

Schwarze, Christoph (2001) On the representation of French and Italian clitics. In: Miriam Butt & Tracy Holloway King (eds.). *Proceedings of LFG01*. Stanford: CSLI Publications. (Available http: csli-publications.stanford.edu/LFG/6/lfg01-toc.html)

Selkirk, Elisabeth (1995) The Prosodic Structure of Function Words. In: Jill N. Beckham, Laura Walsh and Suzanne Urbanczyk (eds.) *Papers in Optimality Theory*. University of Massachsetts Occasional Papers 18, Amherst MA: GSLA, pp. 439-469.

Sells, Peter (1998) Scandinavian Clause Structure and Object Shift. Miriam Butt & Tracy Holloway King (eds.). *Proceedings of LFG98*. Stanford: CSLI Publications. (Availabe http: http://csli-publications.stanford.edu/LFG/3/lfg98.html)

Sells, Peter (2001) *Structure, Alignment and Optimality in Swedish*. Stanford: CSLI Publications.

Sells, Peter (to appear) Syntactic Information and its Morphological Expression. In Louisa Sadler and Andrew Spencer (eds.) *Projecting Morphology*. Stanford: CSLI Publications.

Simpson, Jane and M. Withgott (1986) Pronominal Clitic clusters and Templates. In: Hagit Borer (ed.) *The Syntax of Pronominal Clitics, Syntax and Semantics*, 19, New York/Orlando: Academic Press, pp. 149-174.

Slovacchia, M. et al. (1995) Aspects of Discourse Configurationality in Somali. In: K. Kiss (ed.) *Discourse Configurational Languages*. New York: OUP.

Spencer, Andrew (1992) *Morphological Theory*. Oxford: Blackwell.

Spencer, Andrew (2000) Verbal clitics in Bulgarian: A Paradigm-Function approach. In: Birgit Gerlach and Janet Grijzenhout (eds.) *Clitics in Phonology, Morphology and Syntax*. Amsterdam/Philadelphia: John Benjamins, pp. 355-386.

Spencer, Andrew 2001. "The paradigm-based model of morphosyntax". *Transactions of the Philological Society* 99, 279—313.

Spencer, A. (2003) A realizational approach to case. In: Miriam Butt and Tracy Holloway King (eds.) *Proceedings of the LFG03 Conference*, University at Albany, State University of New York, July 2003. CSLI Publications, http://csli-publications.stanford.edu/

Spencer, Andrew (to appear) Paradigm-based lexicalism: Negation in Japanese. In: Farrell Ackerman, Jim Blevins and Greg Stump (eds.) *Paradigms and Periphrases*. Stanford, CA.: CSLI Publications.

Spencer, Andrew (ms.) Extended Paradigm Function Morphology. Unpublished manuscript, University of Essex. (Available at http://privateweb.essex.ac./~spena/ papers/epfm.pdf)

Stolz, Thomas (1992) Verbale Objectkflexion im Portugiesischen?, *Z. Phon. Sprachwisse. Kommun.forsch.* (*ZPSK*) 45/5, 437-457

Stump, Gregory T. (1991) A Paradigm-based Theory of Morphosemantic Mismatches, *Language* 67, 675-725.

Stump, Gregory T. (1992) On the Theoretical Status of Position Class Restrictions on Inflectional Affixes. In: G. Booij and J. van Marle, eds., *Yearbook of Morphology 1991* Kluwer: Dordrecht, 211-241.

Stump, Gregory T. (1993) Position classes and morphological theory. In: Booij, G. et. al. (eds.) *Yearbook of Morphology 1992*. Kluwer: Dordrecht, pp. 129-180.

Stump, Gregory T. (1997) `Template Morphology and Inflectional Morphology'. In: G. Booij and J. van Marle, eds., *Yearbook of Morphology 1996*. Kluwer: Dordrecht, 217-241.

Stump, Gregory T. (2001) *Inflectional Morphology. A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.

Terzi, Arhonto (1999) Clitic combinations, their hosts and their ordering. *Natural Language and Linguistic Theory* 17: 85-121.

Toivonen, I. (2001) *The Phrase Structure of Non-Projecting words*. PhD Thesis: Stanford University CA.

Vásquez Cuesta, Pilar and Maria Albertina Mendes da Luz. 1971. *Gramática da língua portuguesa*. Lisboa: Edições 70.

Vigário, M. (1998) Elisão da Vogal não-recuada Final e a Palavra Prosódica no Português Europeu. *Actas do XIII Encontro Nacional da APL*. Lisboa: APL/Colibri.

Vigário, M. (1999a) On the Prosodic Status of Stressless Function Words in European Portuguese. In: A. Hall et al. (eds) *Studies on the Phonological Word*. John Benjamins: Amsterdam, pp. 255-294.

Vigário, M. (1999b) Pronominal Cliticization in European Portuguese: a Postlexical Operation. *Catalan Working Papers in Linguistics* 7. Universitat Autònoma de Barcelona: Barcelona, pp. 219-237.

Villalva, Alina (1994) *Estruturas morfológicas. Unidades e hierarquias nas palvras do português*. PhD thesis, University of Lisbon.

Wackernagel, Jakob (1977) Über ein Gesetz der indogermanischen Wortstellung. *Indogermanische Forschungen* I, pp. 333-436.

Wiese, Richard (2000) *The Phonology of German*. Oxford: Oxford University Press.

Yu, A. C. and J. Good (2000) Morphosyntax of Two Turkish Subject Pronominal Paradigms. In: M. Hirotani, A. Coetzee, N. Hall and J-Y. Kim (eds.) *Proceedings of the North East Linguistic Society* 30, pp. 759-773.

Zwicky, Arnold M. (1977) *On Clitics*. Indiana University Linguistics Club.

Zwicky, Arnold M. (1985) Clitics and Particles. *Language* 61.2, pp. 283-305.

Zwicky, Arnold M. (1987) Suppressing the Z's. *Journal of Linguistics* 23.1, pp. 133-148.

Zwicky, Arnold M. (1992) Some choices in the theory of morphology. R. Levine (ed.), *Formal grammar: Theory and implementation.* Oxford Univ. Press 327-71.

Zwicky, A. and G. Pullum (1983) Cliticization vs. Inflection: English n't. *Language* 59 (3), pp. 502-512.