

## Clitics between Syntax and Lexicon

# Linguistik Aktuell/Linguistics Today

*Linguistik Aktuell/Linguistics Today* (LA) provides a platform for original monograph studies into synchronic and diachronic linguistics. Studies in LA confront empirical and theoretical problems as these are currently discussed in syntax, semantics, morphology, phonology, and systematic pragmatics with the aim to establish robust empirical generalizations within a universalistic perspective.

## Series Editor

Werner Abraham

University of Vienna

## Advisory Editorial Board

Guglielmo Cinque (University of Venice)

Günther Grewendorf (J.W. Goethe-University, Frankfurt)

Liliane Haegeman (University of Lille, France)

Hubert Haider (University of Salzburg)

Christer Platzack (University of Lund)

Ian Roberts (Cambridge University)

Ken Safir (Rutgers University, New Brunswick NJ)

Lisa deMena Travis (McGill University)

Sten Vikner (University of Aarhus)

C. Jan-Wouter Zwart (University of Groningen)

## Volume 51

Clitics between Syntax and Lexicon

by Birgit Gerlach

# Clitics between Syntax and Lexicon

Birgit Gerlach

Heinrich-Heine-University

**John Benjamins Publishing Company**  
Amsterdam/Philadelphia



™ The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

## Library of Congress Cataloging-in-Publication Data

Gerlach, Birgit

Clitics between syntax and lexicon / Birgit Gerlach.

p. cm. (Linguistik Aktuell/Linguistics Today, ISSN 0166-0829 ; v. 51)

Includes bibliographical references and index.

1. Grammar, Comparative and general--Clitics. 2. Romance languages--Clitics. I.

Title. II. Linguistik aktuell ; Bd. 51.

P288.G47 2002

415-dc21

2002074771

ISBN 90 272 2772 1 (Eur.) / 1 58811 245 4 (US) (Hb; alk. paper)

© 2002 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands  
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

# Contents

Acknowledgements	IX
Abbreviations	XI
1 Introduction	1
1.1 Definitions of clitics in linguistic research	2
1.2 Problematic clitic phenomena in Romance	5
1.2.1 Clitic paradigms	5
1.2.2 Clitic sequencing	7
1.2.3 Clitic placement and clitic doubling	9
1.3 Theoretical preliminaries	12
1.3.1 Minimalist Morphology	12
1.3.2 Lexical Decomposition Grammar	13
1.3.3 Correspondence Theory	15
2 The status of Romance clitics between words and affixes	17
2.1 Clitic phonology	19
2.2 Clisis vs. inflectional affixation	25
2.2.1 Pronominal clitics	30
2.2.2 Auxiliaries and other clitic-like material	53
2.3 Clisis vs. phrasal affixation	61
2.4 Conclusions	65
3 Clitics in the lexicon	73
3.1 The place of morphology in grammar	74
3.2 Romance clitic inventories	83
3.2.1 Grimshaw's Optimality-theoretical analysis of Italian clitics	86
3.2.2 European and Brazilian Portuguese clitics	89
3.2.3 European and Argentinian Spanish clitics	97
3.2.4 Standard Italian clitics	102
3.2.5 Piattino clitics	110
3.2.6 French clitics	112
3.2.7 Romanian clitics	117
3.3 Conclusions	121

4	Clitic sequences	127
4.1	Properties of clitic sequences in Romance	128
4.1.1	No separation	128
4.1.2	Strict order and combinatory restrictions	129
4.1.3	Opacity	133
4.2	Approaches to clitic sequences in recent research	136
4.3	Interaction of morphology and syntax in clitic sequences	140
4.3.1	The order of clitics	142
4.3.2	Clitic choice when combinations are prohibited	148
4.3.3	Deviations from the clitic order in Romanian	156
4.3.4	Markedness hierarchies and clitic order	159
4.4	Interaction of morphology and phonology in clitic sequences	163
4.4.1	Prosodic minimisation of clitic sequences	163
4.4.2	Dissimilation of clitics within sequences	170
4.5	Conclusions	183
5	Syntax at the interfaces: clitic placement and clitic doubling	189
5.1	Movement and base generation	190
5.2	Clitic placement	195
5.2.1	The domain of cliticisation	198
5.2.2	Clitic split and mesocleft	207
5.2.3	Issues for further research	213
5.3	Clitic doubling	216
5.3.1	Simple cases: prohibited vs. obligatory clitic doubling	220
5.3.2	Person-dependent clitic doubling in Piattino and French	223
5.3.3	Semantic restrictions on clitic doubling	226
5.4	Conclusions	232
6	Summary	237
	References	241
	Appendix	255
I	Paradigms	255
II	Clitic sequence internal phenomena	265
III	Clitic placement and clitic doubling	267
IV	Clitic specific constraints	268

Language Index	271
Name Index	273
Subject Index	275





## Acknowledgements

Several people have assisted me in often surprising ways in the completion of this book and I would like to acknowledge their input here.

First and foremost I would like to thank Dieter Wunderlich who was always willing to share his inexhaustible knowledge of the nature of language. His comments on earlier versions of the book have been extremely supportive. Barbara Stiebels read the pre-final version and I thank her for her careful criticism and detailed comments. For other helpful comments on an earlier version of the book I thank Werner Abraham, James Kilbury and Ana Luis.

I am grateful to Janet Grijzenhout, Sandra Joppen-Hellwig, Ingrid Kaufmann, Renate Lakämper, Albert Ortmann, Alexandra Popescu and Carsten Steins for many fruitful discussions of several topics of this book. Alexandra also provided most of the Romanian data. For assisting me with the French data I thank Anja and Pascal Latrouite and Tahar Guellil. I owe many Spanish data to Teresa Parodi's previous work. For help with the Piattino and Italian data I thank Emanuela Canclini and her family and friends in Piatta. Claudia Kursawe and Stefano Bergonzoni in Bologna also helped me with the Italian data.

The enterprise of writing this book would hardly have been possible without the help of my partner in life Manfred Lebbing and my parents. I dedicate the book to my son Julius, who has been very understanding and diligent during this last year.



# Abbreviations

PRN	pronoun	s	subject clitic	[ ]
PSS	possessive pronoun	O	ambiguous object clitic	[+hr]
P	preposition	DO	direct object clitic	[+hr]
C	complementiser	IO	indirect object clitic	[+hr,+lr]
D	determiner	RFL	reflexive clitic	[+rfl]
ACC	accusative marker	IP	impersonal subject clitic	
DAT	dative marker	PTV	partitive clitic	[+ptv]
		LOC	locative clitic	[+loc]
		NEG	negation clitic	[+neg]
		AUX	auxiliary clitic	[ ]

M	masculine	[-f]	FIN	finite	[+fin]
F	feminine	[+f]	INF	infinite	[-fin]
SG	singular	[-pl]	FUT	future tense	[+fut]
PL	plural	[+pl]	PFV	perfective aspect	[+pfv]
1	1st person	[+1]	SBJ	subjunctive mood	[+sbj]
2	2nd person	[+2]	CND	conditional mood	[+cnd]
3	3rd person	[-1,-2]	IMP	imperative mood	[+imp]
DF	definite	[+df]	GER	gerund	
IDF	indefinite	[-df]	PTC	participle	

MM	Minimalist Morphology
LDG	Lexical Decomposition Grammar
OT	Optimality Theory
CT	Correspondence Theory
HPSG	Head Driven Phrase Structure Grammar

AS	Argentinian Spanish	P	Piattino
ES	European Spanish	F	French
BP	Brazilian Portuguese	I	Italian
EP	European Portuguese	R	Romanian



## CHAPTER 1

# Introduction

This book presents a lexically-based analysis of verbal pronominal clitics in Romance. It is argued that Romance clitics constitute a distinct morphological unit with special clitic-type properties. Therefore, the clitic phenomena are examined from a morphological point of view. Clitics play an increasingly prominent role in other fields of linguistic research, too, and the status of clitics is a topic of much debate in phonology, morphology and syntax, as well as in the respective interfaces. The discussion on clitics has given rise to diverse and controversial analyses. I will look at the phonological, syntactic and prosodic aspects of clitic phenomena insofar as they are relevant to the suggested morphological analysis.

From a morphological point of view, it is questionable whether a distinct morphological category of clitics is linguistically desirable beyond a purely descriptive means. In recent analyses, it has been proposed to accommodate clitics in one of the categories “word” or “affix” (see, e.g., Barbu 1998; Cysmann 2000; Miller 1992; Monachesi 1995; Miller & Sag 1997). Other researchers propose creating a category “phrasal affix” to refer to clitics (see, e.g., Anderson 1992; Halpern 1998; Harrison 1997; Legendre 1996, 2000a,b). These issues will be discussed in chapter 2, taking into account the criteria of Zwicky (1985) and Zwicky & Pullum (1983) as well as Nübling’s (1992) “bundles of scales” to distinguish between affixes, clitics and words. In chapter 3, the relevant clitic paradigms are introduced and the clitics are supplied with lexical entries based on the early-insertion model of *Minimalist Morphology* (Wunderlich & Fabri 1996). In chapter 4, it is shown that the combinatory restrictions that underlie the occurrence of opaque clitics in clitic clusters pose a problem for phonology and syntax and seem to demand a morphological analysis. Chapter 4 provides such an analysis and also looks at the phonological principles that influence the combination of clitics. It will be shown that syntactically-orientated late-insertion models of morphology such as *Distributed Morphology* (Bonet 1995; Halle & Marantz 1994; Harris 1995) or *Paradigm Function Morphology* (Spencer 2000; Stump 1993) cannot satisfactorily

explain the combinatory restrictions observed in connection with clitics. I will argue that the morphological component as proposed in *Minimalist Morphology* (Wunderlich & Fabri 1996) is able to capture the problems specific to clitics, especially since this framework allows for the interaction of phonological, morphological and syntactic constraints in an Optimality-theoretical way (see also e.g., Grimshaw 1997; Legendre 1996). Chapter 5 deals with clitic placement and clitic doubling in Romance languages. It is shown that in recent generative literature, the issue remains unsettled whether these phenomena are best accounted for in an entirely syntactic (Kayne 1975; Sportiche 1996; Uriagereka 1995), a prosodic (Halpern 1992), a prosodic-syntactic (Čavar & Wilder 1994), or an Optimality-theoretical approach (Anderson 1996; Legendre 1996). The morphosyntactic status of clitics is still under debate. One of the main problems is whether clitics are arguments as proposed by Kayne (1975) and many others, or whether they are functional heads as proposed by, e.g., Sportiche (1996), Uriagereka (1995, 2000). The arguments for both positions are considered in chapter 5. The shortcomings of these syntactic analyses, especially when dealing with clitic doubling, will be demonstrated. I will show that Optimality-theoretical analyses are most successful with respect to clitic placement and, in my own analysis, I will adopt the linking model of *Lexical Decomposition Grammar* (Joppen & Wunderlich, 1995).

Before going into the analysis, a definition of clitic-hood is given in section 1.1, and in section 1.2, a short overview of the particular clitic phenomena and clitic-specific problems in the examined languages, i.e. clitic sequencing, clitic placement and clitic doubling. Section 1.3 provides the reader with the basic assumptions of the theoretical framework chosen for the analysis.

## 1.1 Definitions of clitics in linguistic research

A clitic is generally understood to be a word that cannot stand on its own and “leans” on a host word. The term “clitic” originates from the Ancient Greek verb *klineîn* ‘to lean on’ (see Kühner 1834/1966). The main characteristic of clitics is captured by this word: clitics are not able to bear stress or — as in the Ancient Greek case — accent by themselves and therefore have to “lean on” another word — called the “host” — in order to become integrated into a prosodic constituent. In Ancient Greek, as in many other languages, most

clitics are enclitics, i.e. they have to lean on a host to their left. In contrast, proclitics lean on a host to their right. (1a) and (1b) show the Ancient Greek indefinite enclitic pronoun *tis* as a single clitic, while (1c) shows the proclitic group *οὐχ hoos ho en*, consisting of the negation *οὐκ*, the complementiser *hoos*, the relative pronoun *ho* and the preposition *en*.

(1) Ancient Greek clitics

- |    |   |    |                           |
|----|---|----|---------------------------|
| a. | philos    tis   | b. | pólemós tis               |
|    | friend   PRN.IDF  |    | enemy   PRN.IDF           |
|    | 'a friend' or 'any friend'                                      |    | 'an enemy' or 'any enemy' |
| c. | οὐχ   hoos   ho   en   téèi   géei                              |    |                           |
|    | NEG   C       D.M   in   D.DAT.F   earth.DAT.F                  |    |                           |
|    | 'not like the one in the earth'      Kühner (1834/1966: §§ 87f) |    |                           |

As can be seen in (1b), Ancient Greek enclitics may affect the metric structure of their host, since *pólemos* usually bears the main stress on the antepenultimate. The enclitic *tis* causes additional secondary stress on the ultimate.

On the basis of the Ancient Greek data, clitics were originally defined by their phonological defectiveness. Therefore, clitics are a main topic of phonological research and their prosodic structure is under discussion. In addition to their typical lack of stress, clitics are prosodically deficient in that they fail to meet prosodic minimality conditions. Unlike prosodic words, clitics need not consist of a full vowel. Moreover, clitics often exhibit different phonological behaviour from other categories. A central issue in recent phonological theory is how clitics should be prosodified and whether we should distinguish a separate "clitic group" in the prosodic hierarchy, as originally proposed by Nespor & Vogel (1986) and Hayes (1989), or whether this category is superfluous. Under the latter view, which is supported in section 2.1, clitics are attached to — or integrated into — categories such as the prosodic word and the phonological phrase (see e.g., Booij 1996; Selkirk 1995; Zec & Inkelas 1991).

With respect to the category, clitics are usually weak forms of functional elements such as pronouns, determiners, auxiliaries, negation particles and question particles. Zwicky (1977) has introduced the distinction between simple and special clitics. Special clitics differ from simple clitics in that they are not necessarily derived from corresponding full forms in a transparent way and may be placed differently from the corresponding full form if there is one. Nübling (1992) argues that special clitics are at a transitional stage

between simple clitics and affixes. It will be shown that the Romance clitics examined in this study are special clitics. Considering the morphological and syntactic properties of clitics, one can observe that their behaviour in this respect shows some kind of defectiveness as well. A catalogue of criteria to distinguish clitics from affixes on the one hand and to distinguish clitics from words on the other has been drawn up by Zwicky (1985) and Zwicky & Pullum (1983). Nübling (1992) has re-organised the so called Zwicky-criteria in a bundle of scales which also contain some additional criteria concerning the relation of clitics and full forms. This issue will be dealt with in chapter 2.

A comparison of clitics and corresponding full forms is also made by Kayne (1975). He aims at a mainly syntactic description of French, which leads him to a definition of clitics based on their syntactic behaviour compared with full pronouns or NPs. These properties are shared by nearly all Romance pronominal clitics. For Italian and French, consider the sentences in (2a–d) and (2e) from Cardinaletti (1999). The first sentence always contains the full pronoun, i.e. *lui* or *voi* respectively, and the second sentence contains the clitic, i.e. *lo* or *vi* respectively.

(2) Properties of clitics and full pronouns

- |    |     |                                       |                                      |
|----|-----|---------------------------------------|--------------------------------------|
| a. | i.  | Maria conosce solo lui.               | ‘Maria knows only him.’              |
|    | ii. | *Maria <b>lo</b> conosce solo.        |                                      |
| b. | i.  | Maria conosce lui e voi.              | ‘Maria knows him and you.’           |
|    | ii. | *Maria <b>lo</b> e <b>vi</b> conosce. |                                      |
| c. | i.  | Maria conosce LUI, non voi.           | ‘Maria knows HIM, not you.’          |
|    | ii. | *Maria <b>LO</b> conosce, non voi.    |                                      |
| d. | i.  | Chi conosce, Maria? Lui.              | ‘Whom does Maria know?’              |
|    | ii. | *Chi conosce, Maria? <b>Lo</b> .      | ‘Him.’                               |
| e. | i.  | Je ne connais/*achète que lui.        | ‘I don’t know/buy (other) than him.’ |
|    | ii. | Je l’ai connu / acheté.               | ‘I have bought/known him/it.’        |

Contrary to full pronouns, clitics are syntactically deficient in that they cannot be modified (2a), co-ordinated (2b) or contrastively stressed (2c). Furthermore, they can occur neither in isolation (2d) nor in the same positions as full pronouns or NPs. With respect to their referential properties, they differ from



full pronouns insofar as they can refer to non-human entities, while full pronouns cannot (2e).

## 1.2 Problematic clitic phenomena in Romance

Verbal pronominal clitics in the Romance languages are at the centre of this study. A morphological analysis of clitics has to face at least the following three problematic issues. First, it has to be explained that the clitic inventories of the various languages — though phonologically similar — differ a lot with respect to their paradigm structure, i.e. which paradigm cells are filled by distinct elements and which are filled by syncretism. Second, it has to be decided how the distinctive ordering and combinatory restrictions that combinations of clitics with other clitics obey are accommodated in the grammar. Third, the unique conditions on the placement as well as on the occurrence of clitics need a careful re-examination.

Analyses will be given for verbal clitics and verbal clitic phenomena in Standard Italian, Iberian and Rio de la Plata Spanish, Standard French, Iberian and Brazilian Portuguese and Standard Romanian. Furthermore, the northern Italian dialect Piattino spoken in Piatta in Lombardy is analysed. For Piattino, a completely new sample of data is introduced. Other dialects or strata of these languages will be taken into account where necessary or helpful.

### 1.2.1 Clitic paradigms

The first part of the book aims at a description of the inventories of the verbal clitics in the respective languages. As will be seen, all of the languages have pronominal object and reflexive clitics. In addition, we find subject clitics in French and Piattino. Moreover, Italian and French have locative and partitive clitics, and Romanian has clitic auxiliaries. Though subject and object clitics are at the centre of my research, I will also provide lexical entries for the other verbal clitics in the respective languages. Two leading questions concerning the clitic paradigms are examined.

- What are the lexical entries for the respective clitics?
- What kind of syncretism can be observed within the clitic paradigms?

The answers to both questions need a careful exploration of the occurrences of

each clitic in the respective languages. Consider for example the Italian sentences in (3). The Italian clitic *ci* can be used as a 1st person direct object (3a), a 1st person indirect object (3b), a 1st person reflexive pronoun (3c), a prepositional object (3d), a locative pronoun (3e), or an expletive pronoun (3f).

(3) Italian *ci* and its different uses

- |   |                          |
|---|--------------------------|
| a. direct object  | b. indirect object       |
| Arrivederci.  | Ci da questo libro.      |
| see.again.us  | us give.3sg this book    |
| ‘Good bye’  | ‘He gives us this book.’ |
|   |                          |
| c. reflexive  |                          |
| Ci divertiamo.  |                          |
| ourselves enjoy.1PL   |                          |
| ‘We are happy.’   |                          |
|   |                          |
| d. prepositional object   |                          |
| È stata proprio una brutta storia, ma non ci penso più.                 |                          |
| is been really a ugly story but not about.it think yet                  |                          |
| ‘It has really been an ugly story, but I don’t think about it anymore.’ |                          |
|   |                          |
| e. locative   |                          |
| Ci vado domani. Claudia non c’ è.                                       |                          |
| there go.1sg tomorrow Claudia not here is                               |                          |
| ‘I go there tomorrow.’ ‘Claudia is not here.’                           |                          |
|   |                          |
| f. expletive  |                          |
| In Italia ci sono molte grandi città.                                   |                          |
| In Italy there are many big cities                                      |                          |
| ‘There are many big cities in Italy.’                                   |                          |
|   |                          |
| Ce l’ha il biglietto?   |                          |
| there it have.3sg the ticket  |                          |
| ‘Do you (polite) have the ticket?’                                      |                          |

Other clitics such as *gli* or *lo* are also found as determiners. This huge number of different uses can be observed quite often in connection with clitics. With regard to underspecification and syncretism, one needs to decide in each case which different functions should be described by the same entries, i.e. which syncretism should be seen as systematic and which one coincidental. This will be provided by lexical entries in the framework of *Minimalist Morphology* in chapter 3.

### 1.2.2 Clitic sequencing

In Romance languages, clitic combinations maintain a strict internal order, they resist separation, they are often confined to at most two elements, and they exhibit unique morphophonological behaviour. The principles that underlie the building of clitic sequences are the topic of the fourth chapter of this thesis.

In the Italian sentence in (4), the direct object must precede the indirect object if both objects consist of full NPs.

- (4) Order of object DPs in Italian  
 Devo dare il libro a Stefano.  
 must.1SG give the book to Stefano.  
 'I must give the book to Stefano.'

If both objects are replaced by clitics, however, the indirect object clitic *gli* (or its allomorph *glie*, which is restricted to clitic sequences) always must precede the direct object *lo*, regardless of the proclitic (5a) or enclitic (5b) status of the clitic sequence. One of the most noticeable properties of clitic combinations is that the strict order of their elements is maintained with respect to each other and not — as is the case for affixes — with respect to the host. Another property of clitic combinations is that they usually resist separation, even if there are different positions in the sentence which may potentially accommodate clitics (5c,d). In Italian modal verb constructions, for instance, a clitic cluster can attach either to the embedded infinitive as in (5a), or to the finite modal verb as in (5b). Adjacency seems essential and separation of two clitics is ungrammatical (5c,d).

- (5) Order of clitics in Italian
- |   |   |
|---|---|
| a. Devo dar- <b>glie-lo</b> .<br>must.1SG give-him-it<br>'I must give it to him.' | b. <b>Gl</b> ie-lo devo dare.<br>him-it must.1SG give<br>'I must give it to him.' |
| c. * <b>Gl</b> i devo dar-lo.   | d. * <b>Lo</b> devo dar-gli.  |

Moreover, clitics in clitic sequences cannot combine freely. Clitic sequences in Romance languages, for example, are often restricted to two clitics, even if the clitic inventory consists of more than just a direct and an indirect object clitic. Standard Italian, for example, has locative and partitive clitics, and in Standard French we find subject, locative and partitive clitics. Furthermore, re-

strictions on possible combinations of clitics are based on case and person specifications. As mentioned above, an indirect object clitic always precedes a direct object clitic in Italian. In addition, 1st and 2nd person clitics can combine neither with each other nor with a 3rd person indirect object clitic. In these cases, one of the clitics is omitted and the respective argument is expressed by a full DP.

Moreover, 3rd person clitics and reflexive clitics exhibit particular morphological behaviour in clitic sequences (see Bonet 1995:607, who refers to these cases as “opaque clitics”). These clitics are either deleted, or replaced by other clitics, or their surface form in combination with other clitics differs from their form in isolation. The above-mentioned Italian clitic *ci*, for instance, is found as a substitute for the first *si* in combinations of impersonal and reflexive *si* (see 6).

(6) Restriction on the combination of impersonal and reflexive *si* in Italian

- a. I ragazzi *si* rallegrano delle vacanze estive.  
the teens oneself look.forward.3pl to.the holidays summer(adj)  
‘The teens look forward to the summer holidays.’
- b. Quando *si* é ragazzi *ci* *si* / \**si* *si*  
when one is teen *ci* oneself / one oneself  
rallegra delle vacanze estive.  
look.forward.3pl to.the holidays summer(adj)  
‘When one is a teen one looks forward to the summer holidays.’

The following questions shall be examined in the analysis of clitic sequences.

- Which clitic combinations are attested and which are not?
- Are there similar paradigms of clitic combinations in the Romance languages?
- Which are the underlying principles that determine the order of clitics within clitic combinations and the restrictions on possible clitic combinations?
- How do these principles interact?
- When do we find deletion and when do we find substitution within clitic sequences?

We will see that the answers to these questions can be found in the interaction of phonological and morphological constraints.

### 1.2.3 Clitic placement and clitic doubling

The placement of clitics in the sentence is a purely syntactic issue. The following questions need to be answered.

- Why are the positions in which clitics surface different from the positions in which full DPs surface?
- Why are these positions exclusive to clitics?
- How can the co-occurrence of both, clitic and DP, be explained?

To solve these questions, we not only need to consider to which syntactic position clitics are assigned on the surface, i.e. “where do they go”, we also need to ask in which syntactic position clitics are generated, i.e. “where do they come from”. Let us consider the Italian examples in (4) and (5) above again. As illustrated in (5), clitics are adjacent to a verb. Direct object DPs follow the finite verb (see 4), while direct object clitics precede it (see 5b). With regard to morphosyntax, the alternation between pre- and postverbal clitics with, e.g., finite verbs vs. infinitives (see 5a vs. 5b) needs to be explained. Does the clitic-verb complex coalesce into a single morphosyntactic unit or do they both retain a certain independence? For both possibilities, it needs to be explained how the morphosyntactic features of the verb can influence the direction of clisis. A huge number of different syntactic analyses is concerned with these questions for Romance as well as for Slavic languages. Clitic placement in Portuguese, where the morphology/syntax as well as the syntax/semantics interface play a role, is particularly complex. The unmarked position for clitics in European Portuguese is behind the verb including inflectional morphology (see 7a), but in future and conditional tense, clitics occur directly behind the verb stem and in front of the inflectional endings (see 7b). Certain syntactic and semantic triggers (e.g. negation, question, emphasis), however, cause the clitic to occur in front of the verb in all tenses (see 7c).

#### (7) Clitic Placement in European Portuguese

- |                      |                         |    |                   |
|----------------------|-------------------------|----|-------------------|
| a.                   | Eu      vi-te.          | b. | Comprá-lo-ás      |
|                      | PRN1SG see.PST-you      |    | buy-it-FUT.2SG    |
|                      | ‘I saw you’             |    | ‘You will buy it’ |
| c. Não me lembrarei. |                         |    |                   |
|                      | not me remember.FUT.1SG |    |                   |
|                      | ‘I will not remember.’  |    |                   |

In chapter 5, an overview will be given of current syntactic approaches to clitic placement including the two prevailing approaches, namely the “movement approach” (Kayne 1975, among others) and the “base-generation” approach (Jaeggli 1982, among others), and the most recent approaches, which treat clitics as functional D° heads (Sportiche 1996; Uriagereka 1995, 2000). These approaches are compared to recent Optimality-theoretical accounts of clitic placement in Balkan languages (Anderson 1996; Legendre 1996). It is argued that Optimality-theoretical approaches are capable of analysing clitic placement in a particularly elegant way since they allow constraints from different modules to interact. For the Romance languages, an analysis is provided which allows for the interaction of syntactic, morphological and prosodic constraints.

While in Standard French and Standard Italian a clitic and a DP cannot co-occur in the same sentence (see 8a,b), Spanish, Romanian and colloquial varieties and dialects of French and Italian allow clitic doubling, i.e. the co-occurrence of co-referential clitic and DP in the same sentence. Indirect object DPs in, e.g., Spanish are obligatorily doubled by clitics, and direct object DPs may also be doubled by clitics, for instance in Rio de la Plata Spanish and in Romanian (see 8c,d from Jaeggli 1986 and Sportiche 1996 respectively).

(8) The co-occurrence of clitics and DPs in Romance

- |                                 |                     |
|---------------------------------|---------------------|
| a. French                       | b. Italian          |
| Je la vois.                     | La vedo.            |
| Je vois Claudia.                | Vedo Claudia.       |
| *Je la vois Claudia.            | *La vedo Claudia.   |
| I her see.1SG Claudia           | her see.1SG Claudia |
| ‘I see Claudia’                 | ‘I see Claudia.’    |
| c. Rio de la Plata Spanish      |                     |
| Lo veo a Juan.                  |                     |
| him see.1SG to Juan             |                     |
| ‘I have seen Juan.’             |                     |
| d. Romanian                     |                     |
| L-am văzut pe Popescu.          |                     |
| him-AUX.1SG see.PTC ACC Popescu |                     |
| ‘I have seen Popescu.’          |                     |

At first glance, the co-occurrence of clitics and DPs seems to be licensed by the

additional occurrence of a case-marking element (*a* in Spanish and *pe* in Romanian) that precedes the DP. This observation is captured by “Kayne’s Generalisation”. This generalisation, however, does not withstand a closer examination of the data. In Spanish, for instance, the case marker is obligatory for human entities (9a), whereas the clitic is obligatory for specific entities (9b, where *al* stands for the case marker *a* plus the definite article *el*).

(9) The case marker *a* and the clitic in Spanish

- a. \***La** veo      *a* una      mujer.  
       Veo      *a* una      mujer.  
       \***La** veo              una      mujer.  
       her    see.1SG to    D.IDF   woman  
       ‘I see a woman.’
- b. \***Lo**    veo      *al*              libro.  
       \*      Veo      *al*              libro.  
       **Lo**    veo                      *el*      libro.  
       him    see.1SG to.D.DF    D.DF   book  
       ‘I see the book.’

This implies that the presence of the clitic and the presence of the case marker are required (or licensed) by different semantic features of the object. Their co-occurrence in (8c,d) above is coincidental (i.e. dependent on the object having the semantic features [+human] and [+specific]) and not — as predicted by Kayne’s Generalization — necessary.

Recent syntactic approaches are discussed in chapter 5. It is shown in detail that they all fail to explain certain data such as the relevance of person and animacy as a necessary condition for doubling structures. As an alternative, I will suggest that clitic doubling is a kind of agreement relation (in accordance with, e.g., Auger 1993; Suñer 1988; Torrego 1996), restricted by the lexical entries of the clitics on the one hand and (in accordance with Ortmann 2002) grammatical economy principles on the other hand. The requirement of agreement by clitics and the requirement of economy are thus in conflict, a problem which is in turn resolved by an Optimality-theoretical hierarchical ordering of these requirements. Moreover, it is shown that the morphological prohibition of the combination of, for instance, 1st and 2nd person object clitics (see section 1.2.2 above) may also conflict with the agreement requirement. In such cases, the clitics cannot be combined even if both

clitics are required due to their specificity or animacy. Here, it is the requirement for morphological explicitness which governs the choice as to which of the arguments is realised as a clitic and which is realised as a pronoun.

### 1.3. Theoretical preliminaries

In the following, I will briefly outline the main assumptions I will be using in my study, taken from the integrated framework of *Minimalist Morphology*, *Lexical Decomposition Grammar* and *Correspondence Theory* (see also Gerlach 1998a,b; Stiebels 2000a,b; Wunderlich 2001 a,b).

#### 1.3.1 Minimalist Morphology

*Minimalist Morphology* (= MM, Wunderlich & Fabri 1996; Wunderlich 1996a,b,c) as an early-insertion model was originally developed for inflectional morphology. Morphemes are considered to be functors, represented by a lexical entry that consists of a phonological form, several morphosyntactic features, which are understood as output information, and a subcategorisation frame, which is understood as input information. The lexical entries for the single morphemes in the Italian verb form *vedo* ‘I see’, for instance, are represented as follows (see Canclini 1999 for a detailed study of Italian verbal inflection):

(10) Lexical entries

- a. /-o/ : [+min], [+1]/[+V]  $\lambda V \exists x^{<+1>} V(x)$
- b. /ved/ : [ ], [+V],  $\lambda y \lambda x \text{ SEE}(x,y)$
- c. /ved-o/ : [+max], [+1, +V],  $\lambda y \exists x^{<+1>} \text{ SEE}(x,y)$

The inflectional suffix with the phonological form /-o/ has the feature [+1] in the output and the subcategorisation restriction [+V] in the input (10a). Other output features such as [-pl], encoding singular, or [-hr] and [-lr], encoding the grammatical function “subject” (see section 1.3.2 below), are not part of the lexical entry itself but may be added by default. The output of the verb stem /ved/ (10b) consists of its categorical features, i.e. [+V], its semantic form, i.e.  $\text{SEE}(x,y)$ , and its argument structure, i.e.  $\lambda y \lambda x$ . Inflectional affixes are characterised by the feature [+min], indicating that they are morphologically minimal. Words, on the other hand, are morphologically maximal, which is represented by the feature [+max]. Only elements with the latter



feature may be projected into the syntax (see Stiebels & Wunderlich 1994). By combination of the stem and the affix the inflected form *vedo* emerges, which can be projected into the syntax and thus bears the feature [+max]. In the verbal paradigm, this form occupies the cell characterised by the features [+1,+V]. The paradigm is considered to be the interface between syntax and morphology. The subject agreement information is assigned to the respective  $\lambda$ -abstractor, which is changed to the existential quantifier due to the pro-drop properties of Italian, similarly to co-indexing in HPSG (Pollard & Sag 1994). (Note that agreement information in non pro-drop languages such as German, for instance, does not affect the abstractor in such a way.)

The combination of morphemes, i.e. stems and affixes, takes place in the lexicon and follows general principles that may be ordered. MM makes extensive use of underspecification. Only those features necessary for building up the paradigm are specified in the lexical entry. This means that [+]-valued features are part of the lexical entry, while [-] values are added by default. In MM, the following principles are responsible for paradigm building: each cell of a paradigm is occupied (Completeness), more than one element per cell is prohibited (Uniqueness). A form with a more specific output is preferred over a more complex form (Output Specificity), as well as to a more complex form with the same specification (Simplicity).

The structural features for argument roles as well as the linking mechanism are taken from *Lexical Decomposition Grammar*.

### 1.3.2 Lexical Decomposition Grammar

The linking mechanism in *Lexical Decomposition Grammar* (= LDG, Joppen & Wunderlich 1995; Wunderlich 1997) essentially depends on relational features which encode the argument roles of the verb and, at the same time, the linkers, namely morphological case, structural position and agreement morphemes. These features go back to Kiparsky (1989, 1992). They are modified in LDG as follows:

- (11) +(-)hr:    there is a (no) higher role  
       +(-)lr:    there is a (no) lower role

In LDG, the Semantic Form (SF, see 12) determines the argument structure of the verb (TS = theta structure). Each position in the TS is characterised by features depending on whether a lower or higher position exists.

- (12) *give* SF: CAUSE( $x$ , BECOME(POSS( $y$ ,  $z$ )))( $s$ )  
TS:  $\lambda z \lambda y \lambda x \lambda s$

The TS of a (canonical) intransitive verb only contains one argument (in addition to the situation argument, which is ignored here). This means that there is neither a higher nor a lower role. Thus, this position is assigned the features  $[-hr, -lr]$  (13a). With (canonical) transitive verbs, the lowest role (the “direct object”) is assigned  $[-lr]$ , since there is no lower role, and  $[+hr]$ , because of the presence of a higher role (the “subject”). The highest role (the “subject”) is assigned the reverse feature values (13b). For the medial role of (canonical) ditransitive verbs (the “indirect object”), a higher role exists (the “subject”), hence the feature value  $[+hr]$ , as well as a lower role (the “direct object”), hence the feature value  $[+lr]$  (13c).

- (13) Features of the TS positions of canonical verbs
- | a. intransitive        | b. transitive                       | c. ditransitive                                  |
|------------------------|-------------------------------------|--|
| $\lambda x \quad V(x)$ | $\lambda y \lambda x \quad V(x, y)$ | $\lambda z \lambda y \lambda x \quad V(x, y, z)$ |
| $-hr$                  | $+hr \ -hr$                         | $+hr \ +hr \ -hr$                                |
| $-lr$                  | $-lr \ +lr$                         | $-lr \ +lr \ +lr$                                |

The same features define the structural cases, with the difference that they are not fully specified, but only characterised by  $[+]$ -valued features. It is assumed that argument clitics are assigned the same features as structural cases. These case features are part of the clitics’ lexical entries. In nominative-accusative languages, dative case (DAT) and indirect object clitics (IO) are fully specified by  $[+hr, +lr]$ , accusative case and direct object clitics (O) only bear the feature  $[+hr]$ , and nominative case and subject clitics (S) remain unmarked  $[ \quad ]$  and are thus least restricted:

- (14) case features: DAT/IO:  $[+hr, +lr]$  ACC/O:  $[+hr]$  NOM/S:  $[ \quad ]$

Structural linking consists in the choice of the optimal case pattern for a given TS. Ideally, this includes the choice of the most specific compatible linker for each argument role (Specificity Principle) and the use of each linker only once (Uniqueness Constraint). As a result of these principles, the arguments of canonical verbs are linked to structural cases as shown in (15).

## (15) Linking for canonical (nominative-accusative) verbs

a. $\lambda x$	$V(x)$	b. $\lambda y$	$\lambda x$	$V(x,z)$	c. $\lambda z$	$\lambda y$	$\lambda x$	$V(x,y,z)$
-hr		+hr	-hr		+hr	+hr	-hr	
-lr		-lr	+lr		-lr	+lr	+lr	
↓		↓	↓		↓	↓	↓	
NOM		ACC	NOM		ACC	DAT	NOM	
s		O	s		O	IO	s	

The above-mentioned linking principles, however, are not always observed. Restrictions on clitic sequences (see section 1.2.2), for instance, may sometimes force substitution of clitics, so that it is not the most specific clitic that is linked to an argument role. This means that other principles or restrictions may compete with the linking principles and sometimes outrank them.

## 1.3.3 Correspondence Theory

In clitic combinations, it is not only suppletion of clitics which is observed, but also deletion of clitics. Combinations of 1st and 2nd person object clitics, for instance, do not occur, even if both objects should cliticise due to clitic doubling requirements. Moreover, clitics often do not surface in their underlying phonological form. These cases, as will be argued in this study, are the result of the competition of constraints, which may belong to various grammatical modules. It depends on the relative ranking of the constraints whether they must be observed or not. *Optimality Theory* (= OT, McCarthy & Prince 1993; Prince & Smolensky 1993) and its more recent version *Correspondence Theory* (= CT, McCarthy & Prince 1995) provide a suitable mechanism for implementing this idea. In CT, grammar consists of a set of violable ranked constraints. While the constraint ranking is language-specific, the constraints themselves are universal. It is thus possible to account for the differences between languages with the same set of constraints but different rankings. The set of output candidates to a given input is evaluated in view of the constraint ranking. The best (hence optimal) output candidate is the one with the fewest violations of highly ranked constraints. The correspondence relation of input and output candidates is checked by faithfulness constraints (in 16 quoted from McCarthy & Prince 1995).

- (16) Relation of  $S_1$  (input) and  $S_2$  (output) (Faithfulness):

The MAX Constraint Family

Every segment of  $S_1$  has a correspondent in  $S_2$ .

The DEP Constraint Family

Every segment of  $S_2$  has a correspondent in  $S_1$ .

The IDENT(F) Constraint Family

Let  $\alpha$  be a segment in  $S_1$ , and  $\beta$  be any correspondent of  $\alpha$  in  $S_2$ .

If  $\alpha$  is  $[\gamma F]$ , then  $\beta$  is  $[\gamma F]$ .

MAX constraints restrict deletion, while DEP constraints restrict epenthesis. If we generalise the notion of segment, the two constraint families together state that there is a one-to-one correspondence of elements regardless of which elements correspond to each other. IDENT makes sure that corresponding elements are identical with respect to their features. In chapter 3, the input and output and the model of grammar that is suggested here is introduced. For agreement relations, the concept of transitive correspondence is developed. For clitic sequences, a more elaborate formulation of the constraints is needed and, in addition, some morphological and phonological constraints. These constraints will be introduced step by step throughout the analysis in chapter 4. Additionally, alignment constraints are necessary in order to explain the placement of clitics. These constraints are developed in chapter 5. Finally, chapter 5 deals with the constraints restricting clitic doubling.

## CHAPTER 2

# The status of Romance clitics between words and affixes

The morphological status of clitics has been under much discussion in recent research. The debate focuses on whether clitics constitute an autonomous morphological category, or whether they can be described as one of the independently motivated categories “affix” or “word”. A threefold system that distinguishes words, affixes and clitics is argued to be undesirable for reasons of economy. Therefore, many researchers would like to dispense with the term “clitic”, even if useful from a traditional descriptive point of view.

The so-called Zwicky-criteria (Zwicky 1985; Zwicky & Pullum 1983, see below section 2.2) are often used to determine whether a linguistic element is a clitic, an affix, or a word. Nevertheless, it is often difficult to determine the status of these elements, since these criteria are gradual rather than absolute. Therefore, the Zwicky criteria have been supplemented by various other syntactic and morphological criteria (see for instance Kayne 1975; Cardinaletti & Starke 1999; Kaiser 1992; Nübling 1992). Sometimes, an accurate re-examination of the elements that are traditionally called clitics may lead to a re-categorisation of these elements but different researchers may also come to different results. The extraordinary properties of Romance clitics often lead to contradicting analyses. Let us consider in this respect the European Portuguese examples in (1). The pronominal clitics in European Portuguese occur after the verb including its inflectional endings in unmarked sentences (see 1a from Hundertmark-Santos Martins 1998). In future tense and conditional mood, however, the clitics occur before the inflectional endings (see 1b from Hundertmark-Santos Martins 1998). This seems to be clear evidence in favour of the affixal status of these elements. There is, however, also evidence which strongly contradicts the categorisation as affixes. In the presence of certain syntactic and semantic triggers (e.g. negation, question markers, certain adverbs, etc.) the clitics occur preverbally, even in future tense and conditional mood (see 1c).

## (1) European Portuguese clitic placement

- a. Ela tem uma casa linda e ontem  
PRN have.3SG D.IDF house nice and yesterday  
mostrou-no-la.  
show.PST.3SG-10-3F.DO  
'She has a nice house and she showed it to us yesterday.'
- b. Ela tem uma casa muito linda. Mostrar-no-la-á?  
PRN have.3SG D.IDF house very nice. show-10-3F.DO-FUT.3SG  
'She has a very nice house. Will she show it to us?'
- c. Não no-lo mostrar-á.  
not 10-3F.DO show-FUT.3SG  
'She will not show it to us.'

European Portuguese clitics thus display very particular behaviour in that they — like affixes — occur within inflected words but — like words — are affected by their syntactic and semantic environment. I will re-examine the elements called clitics in the Romance languages below with respect to the Zwicky-criteria and show that at least pronominal clitics and auxiliary clitics behave differently from inflectional affixes as well as from words. In section 2.2, I will then argue that the independent category “clitic” is reasonable for more than traditional descriptive grounds.

Another approach which dispenses with the term clitic is Anderson's theory of “phrasal affixation”. Several researchers have adopted Anderson's (1992) suggestion to analyse special clitics as a particular kind of affix. Based on Klavans' (1980, 1985) work on the positional properties of clitics, Anderson argues that clitics behave like affixes and that the only difference is that clitics are adjoined to syntactic phrases, while affixes are adjoined to words. For this reason, Anderson refers to clitics as “phrasal affixes”. I will discuss this proposal in section 2.3.

Whether clitics should be considered an independent category is not only an issue in morphological research but also in phonological research. As already mentioned in the introduction, the original definition of clitics is a phonological one, relying on their prosodic deficiency. Mostly, clitics are prosodically deficient in that they fail to meet prosodic minimality conditions. Unlike prosodic words, clitics need not contain a full vowel. Moreover, clitics often exhibit different phonological behaviour from other categories. In phonology, the prosodic structure of clitics is disputed. A central issue in recent phonological theory is how clitics should be prosodified and whether we