

This is a close-to-final version of
Vigário, Marina (2003) *The Prosodic Word in European Portuguese*. (Interface Explorations Series, 6). Berlin/ New York: Mouton de Gruyter (440 pp) [2nd edition, 2011].
DOI: 10.1515/9783110900927

Please note that the page numbers of the footnotes at the end of each chapter in this document do not match the published version's. In the published version, all notes appear at the end of the book.

Contents

Abbreviations and symbols	xv
Chapter 1	
Theoretical background	1
1. Introduction	1
2. Prosodic Phonology: basic claims	1
3. The organization of grammar	10
3.1. The lexical/postlexical distinction and the properties of phonological processes	10
3.2. The locus of construction of prosodic domains	14
4. The status of the Clitic Group within the prosodic hierarchy	17
5. The Strict Layer Hypothesis	19
6. The prosodization of words	22
6.1. Diagnostics for the prosodic word	22
6.2. The prosodic word domain	24
6.3. The prosodization of clitics	26
6.4. The prosodization of compounds	31
6.5. The prosodization of derived words	34
7. Syllabification, resyllabification, and prosodic restructuring	35
8. Concluding remarks and outline of the book	37
Chapter 2	
Previous studies on	
European Portuguese word phonology	41
1. Introduction	41
2. Pregenerative studies	41
3. SPE studies	46
4. Post-SPE studies	48
4.1. Words with two <i>primary</i> stresses	48
4.2. Cliticization	53
4.3. Phonological processes and the organization of grammar	59
5. Concluding remarks	60

Chapter 3	
Phonological phenomena: description and typology	63
1. Introduction	63
2. Word stress	64
3. Vowel reduction	67
4. Theme vowel deletion	73
5. Final nasal diphthongization	74
6. /e/-centralization	78
7. Glide insertion to break a hiatus	83
8. Lowering of stressless vowels in final syllables closed by /r/	85
9. Initial /r/-strengthening	89
10. The realization of word initial vowels	92
11. Semivocalization	99
11.1. V2 semivocalization (>VG)	100
11.2. V1 semivocalization (>GV)	102
12. Final non-back vowel deletion	104
13. Final round vowel deletion	108
14. Final central vowel deletion	112
15. Syllable degemination	114
16. Initial stress	118
17. Emphatic stress	120
18. Tonal association	121
19. Focal stress	123
20. Summary	124
Chapter 4	
On the affix/clitic status of stressless pronouns	127
1. Introduction	127
2. Potential evidence for the lexical attachment of pronominal clitics	128
2.1. Specialized host	129
2.2. Phonological idiosyncrasy	129
2.3. “Inflection” after cliticization	130
3. Evidence for the postlexical insertion of pronominal clitics	131
3.1. Distributional facts	131
3.2. Phonological facts	134
3.3. Other facts	137

4. On the marks of lexicalization	139
4.1. Selectivity with respect to the host	139
4.2. Phonological idiosyncrasy	141
4.2.1. On the nature of the formal alternations affecting verbs and clitics	141
4.2.2. Towards an account of the phonological idiosyncrasy in verb-clitic sequences	144
4.3. Mesocclisis	147
5. Conclusion	155

Chapter 5

The Prosodic Word 157

1. Introduction	157
2. Phonological properties of prosodic words	157
3. Derived and inflected words corresponding to a single stress domain	163
3.1. Prosodization of suffixes	163
3.2. Prosodization of prefixes	165
4. Host plus clitic combinations	173
4.1. Phonological characterization of clitic words	173
4.1.1. On the definition of clitics	173
4.1.2. Identification and classification of EP clitic words	175
4.2. Lexical prosodization of clitics	182
4.3. Postlexical prosodization of clitics	184
4.3.1. Prosodization of enclitics	186
4.3.2. Prosodization of proclitics	195
5. Affixes versus clitics	204
5.1. Suffixes and enclitics	204
5.2. Prefixes and proclitics	205
6. Conclusion: towards the definition of the prosodic word domain in EP	208

Chapter 6

Compound Prosodic Words 215

1. Introduction	215
2. Phonological characterization of prosodic words and their concatenation	216
3. Prosodization of words with more than one stress domain	218
3.1. Suffixed words with more than one word stress	219

3.2. Words with stressed prefixes	228
3.3. Morphological compounds (stem compounds)	233
3.4. Syntactic compounds (word compounds)	236
3.5. Abbreviations	241
3.6. Mesoclititic structures	243
3.7. Specific combinations of words	247
3.8. Interim summary	250
4. Additional evidence for the prosodic word domain	250
5. Discussion	255
5.1. On the sources of word stress	255
5.2. On the relation between word stress and the prosodic word	260
5.3. Definition of the prosodic word domain in EP: final remarks	262
Chapter 7	
On the reduction of clitics	273
1. Introduction	273
2. EP vowel reduction and other reduction phenomena	278
3. Data collection: materials and procedure	281
4. Results	283
4.1. Consonant-schwa clitics	283
4.2. Other clitics: <i>para</i> , <i>pelo/pela</i> , <i>ao</i> and <i>em</i>	291
4.3. The clitic <i>com</i>	295
4.4. Summary	297
5. Discussion	298
5.1. The phonetic shape of clitics and EP general processes of reduction	298
5.2. Reduction phenomena and word frequency	303
5.3. Generalizations on the reduction of very frequent words	307
5.4. Factors (dis)favoring reduction	309
5.4.1. The importance of intonational phrase initial position	309
5.4.2. The importance of grammatical category	314
5.5. Reduction versus allomorphy	319
6. Summary of main findings	323
Chapter 8	
Conclusion and directions of future research	325
1. Introduction	325
2. Minimal prosodic words	325

3. Clitics and their postlexical attachment	326
4. Compound prosodic words	329
5. Our findings and phonological theory	331
6. Our findings and EP phonology	335
7. Our findings and language acquisition	339
8. Final word	341
Appendix I	343
Appendix II	347
Notes	353
References	407
Index	433

Abbreviations and symbols

#: morphological word boundary
.(dot): syllable division
+: morpheme boundary
1/2/3Pl: 1st/2nd/3rd person plural
1/2/3Sg: 1st/2nd/3rd person singular
A: adjective
ACC: accusative pronominal clitic
ADV: adverb
ART: definite article
AUG: augmentative suffix
CG: clitic group
CL: clitic
COND: conditional
DAT: dative pronominal clitic
DIM: diminutive suffix
EP: European Portuguese
FEM: feminine
FUT: future
I: intonational phrase
IMP: indicative past imperfect
IMPS: impersonal pronominal clitic
INF: infinitive
Lex: lexical syntactic head
MASC: masculine
N: noun
NP: nominal phrase
PAST: indicative past perfect
PB: Brazilian Portuguese
PL: plural
PN: person/number suffix
PP: prepositional phrase
PRES: indicative present
REF: reflexive pronominal clitic
SD: syllable degemination
SG: singular

xvi *Abbreviations and symbols*

SLH: Strict Layer Hypothesis

SUBJ IMP: subjunctive imperfect

SUBJ: subjunctive present

T*: pitch accent

TMA: tense/mood/aspect suffix

TV: theme vowel

U: utterance

V: verb

V1: first vowel

V2: second vowel

VP: verbal phrase

VR: vowel reduction

W1: first word

W2: second word

^x: used in acceptability judgments, denotes a very marked almost impossible realization

Σ: foot

φ: phonological phrase

σ: syllable

ω: prosodic word

Chapter 1

Theoretical background

1. Introduction

This book investigates the prosodic word in European Portuguese (EP). The account of the phonology of various types of words developed here is couched within the Prosodic Phonology framework (e.g. Selkirk 1984; Nespor and Vogel 1986; Hayes 1989). It crucially presupposes, in addition, that the grammar is composed of two distinct levels – the lexical and the postlexical level – as also assumed, in particular, in the Lexical Phonology framework (Kiparsky 1982, 1985; Mohanan 1986; Booij and Rubach 1987; Booij 1994, among others).

In this chapter we present the basic claims of prosodic theory (section 2), as well as the general assumptions concerning the organization of grammar assumed in this book (section 3). We review, furthermore, some of the most relevant proposals found in the literature on the issues investigated in this work (sections 4 to 7). We conclude the chapter with an overview of the book.

2. Prosodic Phonology: basic claims

Among the fundamental pioneering work on prosodic phonology are Selkirk (1980, 1984, 1986), Nespor and Vogel (1982, 1983, 1986), and Hayes (1989). In the next paragraphs we survey some of the basic tenets of the theory, following in particular the view presented in Nespor and Vogel's work.

The issues more specifically related to the prosodic word domain – the constituent studied in this book – are the subject matter of sections 4 to 7. It is only then that we review systematically a small part of the impressive amount of research in prosodic phonology that has followed the early works cited in the preceding paragraph.

Whenever possible, we enrich the selected exemplification with data extracted from investigations already conducted on the prosodic phonology of EP.

2 Theoretical background

§1. Prosodic phonology is primarily a *theory of prosodic structure* – a structure built with reference to morphosyntactic structures, but distinct from them – which is taken to define the domain of pure phonological phenomena.¹

The view that phonological processes refer to phonological constituents, rather than to constituents morphosyntactically defined, emerges from the observation that pure phonological rules systematically fail to see certain information that is present in morphosyntactic representations: for example, they do not refer to category labels, such as N, V or A, and they are not sensitive to empty syntactic elements. In addition, pure phonological rules apply within domains that do not always coincide with the domains defined by syntax or morphology (see also paragraph 7 below).

§2. *Prosodic hierarchy is assumed to be universal*, and to include the constituents listed in (1).² Thus, in all languages of the world it is expected that each of these phonological constituents – and only these – play some role in the phonology of the language.³

- (1) Utterance (U)
- Intonational Phrase (I)
- Phonological Phrase (ϕ)
- Prosodic Word (ω)
- Foot (Σ)
- Syllable (σ)

In addition to this set of prosodic constituents, Hayes (1989) proposes the existence of another domain, located between the prosodic (or phonological) word and the phonological phrase – the clitic group (CG). This proposal has been adopted in Nespor and Vogel (1986) and in much subsequent work. However, more recent investigation has brought up a number of arguments against the clitic group as a prosodic domain, which we review in section 4.⁴

Although prosodic hierarchy is claimed to be universal, it is often observed that some prosodic domains seem to be absent in some languages. Nevertheless, as pointed out in Nespor and Vogel (1986: 11–12), the lack of rules referring to a particular prosodic domain in a given language does not necessarily mean that that prosodic domain does not play a role in the phonology of the language.

A good example of such a case is provided by the phonological phrase in European Portuguese (EP): while no sandhi phenomena have been found

so far to apply with reference to this prosodic domain, thus suggesting the absence of the ϕ -domain in this language, there are other (more subtle) phonological phenomena that show the relevance of this constituent in EP phonology. These phenomena include, (i) stress clash resolution processes – which reflect ϕ -level prominence relations, and hence the phonological phrase domain (cf. Frota 2000: chap.3); (ii) pitch accent distribution – since pitch accents in EP are primarily assigned to the head of a ϕ -phrase (cf. Frota 2000: 4.2.2; Vigário 1998a); (iii) register shift – for register shift only occurs across ϕ s, but not within ϕ s (cf. Vigário 1997a, 1997b, 1998a: 6.2.3.4); and (iv) phonological weight requirements on certain syntactic constructions – since, for example, the I-phrase that includes the clause from which a topicalized phrase is extracted requires a heavy head, that is, a ϕ that either bears focus or is branching (cf. Frota and Vigário 1996, 2002; Vigário and Frota 1998).⁵

§3. The constituents in (1) are arranged in a hierarchy argued to satisfy the *Strict Layer Hypothesis* (Selkirk 1984), formulated in Nespor and Vogel (1986: 7) in the following terms:⁶

(2) Strict Layer Hypothesis (SLH)

- i. A given nonterminal unit of the prosodic hierarchy, X^p , is composed of one or more units of the immediately lower category X^{p-1} .
- ii. A unit of a given level of the hierarchy is exhaustively contained in the superordinate unit of which it is a part.
- iii. The hierarchical structures of prosodic phonology are n-ary branching.
- iv. The relative prominence relation defined for sister nodes is such that one node is assigned the value strong (s) and all other nodes are assigned the value weak (w).

In recent years, a number of studies have shown the need of relaxing some aspects of the Strict Layer Hypothesis. In fact, in order to account for certain phonological phenomena it has been pointed out that it is necessary to allow, for example, for adjunction and compound structures and such prosodic configurations induce to SLH violations (e.g. Ladd 1992; Selkirk 1996; Booij 1996a; Peperkamp 1997a – see also section 5).⁷

4 Theoretical background

§4. There is *no single definition for each prosodic domain*, first of all because the information included in the algorithms of construction of prosodic domains varies to a certain extent across languages. In order to illustrate the type of information required in the syntax-phonology mapping, we present below the definitions of the higher level prosodic domains: the U-domain, as proposed by Nespor and Vogel (1986), and the ϕ and I-domains, as proposed for EP in Frota (2000) (the prosodic word will be dealt with in section 6).

The utterance, the highest constituent of the prosodic hierarchy, is defined in Nespor and Vogel (1986: 222) as a prosodic constituent formed by one or more intonational phrases which includes all the constituents dominated by the highest syntactic node (see 3).

(3) Phonological Utterance

- a. U domain: the domain of U consists of all the Is corresponding to X^n in the syntactic tree.
- b. U construction: Join into an n-ary branching U all Is included in a string delimited by the definition of the domain of U.

Examples of phonological utterances, based on Nespor and Vogel (1986: 8.1), are provided in (4).

- (4) a. *My cousin collects snakes. Gertrude prefers butterflies.*
[[My cousin collects snakes]_I]_U [[Gertrude prefers butterflies]_I]_U
- b. *Our next door neighbor, Mr. Jones, bought an ocelot last week.*
[[Our next door neighbor]_I [Mr. Jones]_I [bought an ocelot last week]_I]_U

The intonational phrase in EP is defined in Frota (2000: 365) as in (5), which is adapted from Nespor and Vogel's (1986) proposal on the I-domain formation.

(5) Intonational Phrase (I) Formation (EP)

- a. I-domain: the domain of I-formation may consist of
 - i. all the ϕ s in a string that is not structurally attached to the sentence tree,
 - or ii. any remaining sequence of adjacent ϕ s in a root sentence.
- b. I-construction: the constituents included in an I must bear a head/complement relation.

Frota includes in a separate condition, stated in (6), a weight requirement to which EP intonational phrases are subject.

- (6) Weight conditions on Is (EP): long phrases tend to be divided; balanced phrases, or the longest phrase in the rightmost position, are preferred.

Illustrative examples of different types of intonational phrases are provided in (7) (from Frota 2000).

- (7) a. [As alunas]_I [até onde sabemos]_I [obtiveram boas avaliações]_I
 ‘The students, as far as we know, have got good marks.’
 b. [O nível actual da inflação]_I [é um indicador económico positivo]_I
 ‘The present level of inflation is a good economic index.’

The phonological phrase, in turn, is defined as in (8) (Frota 2000: 365, adapted from Nespor and Vogel 1986 and Hayes 1989).⁸

- (8) Phonological Phrase (ϕ) Formation (EP)
- a. ϕ -domain: The domain of ϕ -formation is defined by the configuration [... Lex XP ...]_{Lex^{max}};
 - b. ϕ -construction: Elements around Lex are organized into ϕ s so that
 - i. all elements on the non-recursive side of Lex which are still within Lex^{max} are contained in the same ϕ with Lex;
 - ii. a ϕ may optionally contain (i) and a following phrase in the domain of (a).

ϕ -formation is also claimed to be subject to a (minimality) condition, which is stated in (9).⁹

- (9) Branchingness (or weight) condition on ϕ s (EP): a ϕ should contain more material than one prosodic word.

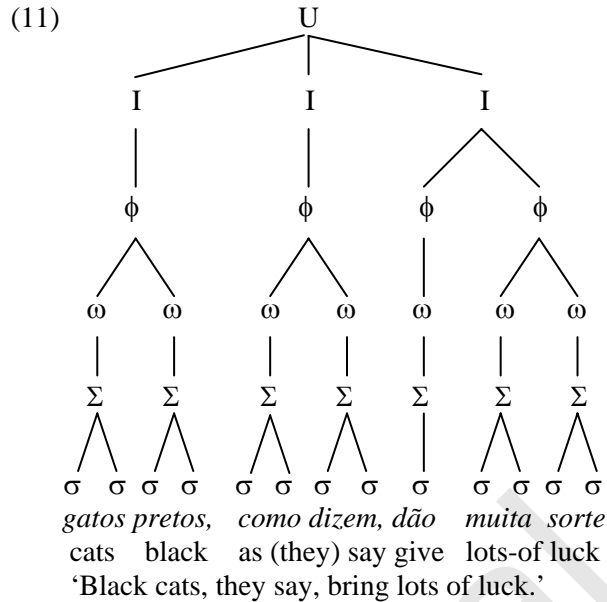
ϕ -phrasing is illustrated in (10) (adapted from Frota 2000).¹⁰

- (10) a. [O bailarino] _{ϕ} [anda sempre] _{ϕ} [de limusine preta] _{ϕ}
 ‘The dancer always drives a black limousine.’

6 Theoretical background

- b. [As americanas]_φ [ofereceram]_φ [a enciclopédia]_φ [ao jornalista]_φ
 ‘The Americans gave the encyclopedia to the journalist.’

The complete tree in (11) further exemplifies how the elements in the segmental string of an EP sentence are grouped into prosodic domains.¹¹



§5. Although the prosodic domains are taken to be *universal*, the information that is included in each domain construction may vary, since it is partially defined on a *language particular basis*.

Thus, for example, Vogel and Kenesi (1987, 1990) argue that focus information is included in the algorithm of the intonational phrase construction in Hungarian, and thus focus affects prosodic structure. In languages like EP, by contrast, focus information seems not to affect prosodic phrasing. This is empirically shown in Frota (2000), who claims that focus is not part of the syntactic information that plays a role in the mapping algorithms of this language.¹²

§6. As mentioned above, it is assumed that *different types of phonological phenomena* refer to prosodic hierarchy. These include not only segmental processes, but also durational, prominence, rhythmic and intonational

phenomena. Examples of each type of phenomenon that have been argued to occur within or at either edge of a prosodic domain are provided below.¹³

Segmental processes that have been shown to refer to prosodic constituents include *t*-flapping in North American English, which is bound by the utterance domain (e.g. Nespor and Vogel 1986); central vowel merger which is bound by the intonational phrase in EP (Frota 2000); *Raddoppiamento Sintattico* which is a prosodic word juncture rule that applies within the phonological phrase domain in Italian (Nespor and Vogel 1986); cluster reduction in Sanskrit which applies at the right edge of the phonological word (Selkirk 1980); *t*-aspiration in English which applies at the left-edge of a foot (e.g. Nespor and Vogel 1986; Selkirk 1996); and *l*-velarization in EP which applies at the right-edge of the syllable domain (e.g. Mateus and d'Andrade 2000).

Prosodic constituents such as the intonational phrase and the utterance constitute domains for *durational phenomena*, since these constituents are often reported to induce preboundary lengthening and their limits identify the locus for pause insertion (e.g. Nespor and Vogel 1986, and Cambier-Langeveld 1997, 2000 for Dutch, and Frota 2000 for EP).

Within prosodic phonology, all non-terminal prosodic constituents are assumed to have a single strong element – the head of the constituent – while all the sister nodes of the head are weak (Nespor and Vogel 1986). The prosodic constituents hence form domains for *prominence assignment*. In EP, like in other right-branching languages (cf. Nespor and Vogel 1986), the head of the ϕ -domain is the rightmost daughter of ϕ . The head of higher-level prosodic constituents is also rightmost (Frota 1997, 2000; Vigário 1998a).

The constituents of the prosodic hierarchy also define the domains of *secondary prominence assignment* and *rhythmic phenomena*. For example, according to Roca (1986), the left edge of the ϕ -phrase defines the locus of occurrence of an optional initial stress in Spanish; Stress Retraction in Italian is claimed in Nespor and Vogel (1986, 1989) to apply between words that belong to the same ϕ ; and, like in many other languages, ϕ -level prominence seems to play a role in the definition of clashing sequences in EP (cf. Frota 1995, 2000: chap.3).

Finally, *intonational phenomena* refer to prosodic structure as well. For example, tune association in Bengali is shown in Hayes and Lahiri (1991) to be accomplished with reference to the prosodic hierarchy; and the same approach has been proposed for EP in Frota (1994a, 2000), Falé (1995) and Vigário (1998a).

8 *Theoretical background*

From this brief illustration, it is clear that the prosodic hierarchy is more than a structure of domains for sandhi or juncture processes. We should notice, nevertheless, that it is still a matter of controversy whether the structure relevant for the different kinds of phonological phenomena is one and the same. For example, while Selkirk (1984: 30) proposes that the domains for segmental processes and for prominence relations coincide, and Hayes and Lahiri (1991) argue in favor of the existence of a single structure to account for segmental and tonal phenomena in Bengali, Gussenhoven and Rietvelt (1992) hold that durational and intonational phenomena in English refer to distinct prosodic hierarchies. In other cases, as in Pierrehumbert and Beckman (1988), the structure assumed to account for intonational phenomena is very similar to the prosodic hierarchy of Selkirk (1984) and Nespor and Vogel (1986), but it is not clear whether the two structures are in fact the same.

For EP, Frota (2000) has extensively shown that various phonological phenomena refer to the same prosodic structure: namely, sandhi phenomena, clash resolution, preboundary lengthening, pause insertion, and intonational phenomena.

§7. A basic claim of prosodic phonology is that the domains of pure phonological phenomena are *not necessarily isomorphic* to morphological or syntactic domains. A classical example of mismatch between the two structures is provided in (12) (from an original pair presented in Chomsky and Halle 1968, cited in Nespor and Vogel 1983, 1986, Hayes 1989, Drescher 1996, among many others).

(12) *Syntactic bracketing*

This is [_{NP} the cat [_S that caught [_{NP} the rat [_S that stole [_{NP} the cheese]]]]]

Prosodic bracketing (at the level of the intonational phrase)

[This is the cat]_I [that caught the rat]_I [that stole the cheese]_I

Mismatches between morphosyntactic structure and prosodic structure may result from the mapping algorithms that relate the two structures, as in the example above. An example of a mismatch at the level of the ϕ -phrase in EP is further shown in (13): if a phonological phrase is defined as in (8), and is subject to the condition formulated in (9) above, it includes the head of a syntactic constituent and it may comprise a following syntactic phrase

within that constituent as well; thus, in a sentence as the one in (13), there is no syntactic constituent that coincides with the second ϕ -phrase.

(13) *O João deu rosas à Maria.*

‘John gave roses to Mary.’

- | | |
|---|----------------------|
| a. [[O João] _{NP} [deu [rosas] _{NP} [à Maria] _{PP}] _{VP}] | syntactic bracketing |
| b. [[O João] _{ϕ} [deu rosas] _{ϕ} [à Maria] _{ϕ}] | prosodic bracketing |

Other possible sources of mismatches between morphosyntactic structures and phonological structures arise from the more flexible nature of the latter. For example, while the syntactic structure of the sentence in (14) is the one represented in (14a), the phonological structure assigned to it may vary depending on factors such as speech rate, or the length of the prosodic constituents (see 14b–d, from Nespor and Vogel 1983).

(14) *The frog ate a fly for lunch.*

- | | |
|--|------------------------|
| a. [[the frog] _{NP} [ate [a fly] _{NP} [for lunch] _{PP}] _{VP}] | syntactic bracketing |
| b. [the frog] _i [ate a fly] _i [for lunch] _i | } prosodic bracketings |
| c. [the frog] _i [ate a fly for lunch] _i | |
| d. [the frog ate a fly for lunch] _i | |

An interesting consequence of the non-isomorphism between syntactic structures and prosodic structures is that a sentence may be non-ambiguous from a syntactic point of view, but ambiguous from a prosodic point of view, as the sentence in (15) (based on Nespor and Vogel 1986: 261–262).

(15) *Marco ha guardato la ragazza col canocchiale.*

‘Marco looked at the girl with the binoculars.’

Syntactic structure

[[Marco]_{NP} [ha guardato [la ragazza]_{NP} [col canocchiale]_{PP}]_{VP}]

‘Marco is holding the binoculars.’

[[Marco]_{NP} [ha guardato [[la ragazza [col canocchiale]_{PP}]_{NP}]_{VP}]

‘The girl is holding the binoculars.’

Prosodic structure (for both interpretations)

[[Marco] _{ϕ} [ha guardato] _{ϕ} [la ragazza] _{ϕ} [col canocchiale] _{ϕ}]_I

§8. Ever since the early proposals on prosodic phonology, a clear distinction is made between *pure phonological processes* and *lexical processes* (e.g. Nespor and Vogel 1986: 2.1.1; Hayes 1990): the former apply auto-

10 *Theoretical background*

matically, postlexically, and refer only to phonological information, including the prosodic domains; the latter apply within the lexicon and may refer to non-phonological information.

In addition, at a first moment it was thought that prosodic domains were built postlexically. Later research, however, in particular by Booij (1988), Inkelas (1990), and Booij and Lieber (1993), led to a refinement of the original idea, so that the constituents up to the word level are now usually assumed to be already present within the lexical component. These issues, which bear on the organization of grammar, are developed in the following section.

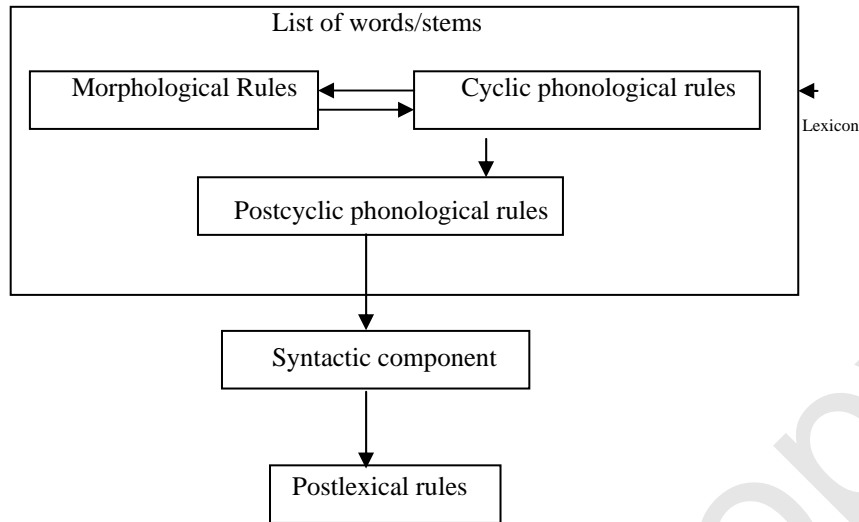
3. The organization of grammar

3.1. The lexical/postlexical distinction and the properties of phonological processes

As mentioned in paragraph 8, standard proposals in prosodic phonology adopt the lexical phonology conception that the grammar (and phonology) is composed of two basic components: the lexical level and the postlexical level (e.g. Nespor and Vogel 1986: chap.1).¹⁴ Regardless of further subdivisions within the lexical level¹⁵, this organization allows us to establish two major classes of phonological processes, according to the point in the grammar where they operate, each of them with well identified properties, as we will see below.

Within the classical lexical phonology framework, both morphology and (some) phonology are assumed to operate in the lexical level, and to be interrelated so that the result of a morphological operation may provide the input for the application of a phonological rule, and conversely. Furthermore, morphological operations may be sensitive to phonological information, and phonology may refer to the morphological structure of words, in addition to idiosyncratic information of specific items or classes of items. This type of interaction between morphology and phonology is claimed in Booij and Rubach (1987) to occur before the word-level, where rules are considered to be cyclic (that is, to reapply after each word-formation operation). At the word-level, on the contrary, phonological rules no longer interact with morphology, and apply both inside and across morpheme boundaries. This conception is represented by the model in (16) (taken from Booij and Rubach 1987: 3).

(16)



In this section we do not intend to provide a full description of the basic claims of lexical phonology (see, for example, Kiparsky 1982, 1985; Mohanan 1982, 1986; Booij and Rubach 1984, 1987; and, for reviews of the theory, Kaisse and Shaw 1985, and, more recently, Booij 1994, 1996b). Rather, we will concentrate only on those aspects that are of crucial importance for the purposes of this book: namely, the properties of the phonological processes that apply in the lexical component, on the one hand, and in the postlexical component, on the other hand.

Of the works mentioned above we extracted the major properties commonly attributed to the lexical rules, which are summarized in (17).

(17) *Lexical Rules*

- only lexical rules may refer to morphological information – thus, for example, (i) phonological rules that distinguish derived environments from underived environments (those that are subject to the Strict Cycle Condition – cf. Kiparsky 1982), necessarily apply within the lexical component; (ii) phonological rules that apply only with specific morphemes or word classes, necessarily apply in the lexical component
- only lexical rules may have exceptions – thus, if a rule does not apply in all cases where its structural description is met, the rule necessarily applies in the lexical component

12 *Theoretical background*

- only lexical rules may precede other lexical rules – thus, if a rule R_α has some property of a lexical rule, a rule R_β that applies prior to R_α necessarily belongs to the lexicon¹⁶
- lexical rules are necessarily categorical, that is they do not have gradient outputs

An additional property often ascribed to lexical rules is structure preservation (e.g. Kiparsky 1985: 92; Kaisse 1990). That is, lexical rules may build structure and specify underspecified segments, but they cannot modify existing structure or the underlying feature composition of segments. Thus, for example, according to this principle the application of a lexical process should not create segments that do not belong to the underlying inventory of the language. However, this property seems to characterize only cyclic rules, as suggested in Booij and Rubach (1987). In fact, these authors show that word-level (i.e. postcyclic, but lexical) processes do create segments that do not exist in the underlying inventory of languages. Canadian French provides an illustrative example. In this language, high vowels alternate between [+tense] in open syllables and [-tense] in closed syllables. Lax vowels do not exist in the underlying system. Nevertheless, the rule that creates lax vowels in closed syllables must operate within the lexical level, since it is not bled by resyllabification. The latter process applies postlexically across words, causing the lexical closed syllables to become open. According to Booij and Rubach, resyllabification does not affect the application of the laxing rule, because it has already applied within the lexical level (see Löfstedt 1992 and Booij 1994: 4.2, among others, for further exemplification of word-level processes that are not structure preserving). Thus, as a general statement, it is not the case that all lexical rules are structure preserving.¹⁷

It is frequently assumed, further, that rules that apply across words operate necessarily in the postlexical component (e.g. Kiparsky 1982). This idea is challenged, for instance, by Hayes (1990), who observes that certain processes that apply between words may have the properties of lexical processes, and thus should belong to the lexical component. In order to account for the existence of such processes in languages like Ewe, French, Kimatumbi, Mende and Hausa, Hayes proposes that these rules are precompiled in the lexical level.

An example of a precompiled rule, taken from Hayes (1990), is tone raising in Ewe. In this language, a High tone in a verb becomes a Rising tone if it follows a High or a Rising tone, and it is followed by a *noun*

starting with a Mid or a Low tone (cf. Clements 1978). This rule is formalized by Hayes as in (18).

- (18) $H \rightarrow R / [_{\text{v}} [-\text{Lopitch}] _] [_{\text{Frame1}}]$
 Frame 1: / $_ [_{\text{N}} [-\text{Hipitch}] \dots]$

This means that the form to be inserted in the context defined by the Frame 1 will have undergone H-tone raising rule in the lexical component.

Precompilation may also consist of lexical listing when the variations in the form of morphemes do not follow from phonological rules. For example, in order to account for the English alternations in the determiner *a/an*, Hayes proposes that the lexical entry of the indefinite article includes the two forms, together with the contexts where they are inserted, as shown in (19).¹⁸

- (19) $\left(\begin{array}{c} \dots \\ \text{Phonological instantiation:} \\ /ə\text{n}/ \text{ (in the context) } / _ \text{ V} \\ /ə/ \text{ (elsewhere)} \end{array} \right)$

Both in Kaisse (1990) and in Hayes (1990) we find a number of diagnostics for lexical-like processes that apply between words. The major properties associated with the application of this type of processes, that is of precompiled phrasal rules in Hayes' terms, as opposed to pure phonological rules applying across words, are listed in (20).¹⁹

- (20) *Precompiled Rules* (lexical rules applying across words)²⁰
- only precompiled rules can precede rules of lexical phonology
 - only precompiled rules can precede morphological rules
 - precompiled rules cannot follow postlexical rules
 - only precompiled rules can treat parallel X' categories asymmetrically, that is, they may refer to properties of individual items or classes of items
 - only precompiled rules may show sensitivity to syntactic information, including elements with no phonetic content (empty categories, traces, pro)
 - precompiled rules are necessarily ingradient

14 *Theoretical background*

As for the properties of postlexical processes, they are listed in (21), taken from various sources, such as Kaisse and Shaw (1985), Nespor and Vogel (1986), Hayes (1990), Kaisse (1990), and Nespor (1990).

(21) *Postlexical Rules*

- postlexical rules do not refer directly to morphosyntactic or lexical information, and thus their application is general and free of idiosyncratic exceptions (in other words, postlexical rules apply automatically, and refer only to phonological information, including the prosodic domain within which they operate)
- postlexical rules always follow lexical (including precompiled) rules
- the output of a postlexical rule may be gradient
- postlexical rules are often optional
- postlexical rules may be sensitive to speech rate²¹

Notice that, as we have seen above, the fact that postlexical rules apply to strings that may be larger than the word is not a distinctive property of postlexical rules, since there are also lexical rules applying across words. Similarly, the absence of structure preservation is not specific of postlexical rules, and thus it is not included among the properties that may identify these rules.²²

Mohanan (1982) provides, in addition, evidence for the psychological reality of the lexical/postlexical levels (cf. Kaisse and Shaw 1985): speakers judge the output of lexical rules as different from the input, while this is not always the case of postlexical processes; the forms that result from slips of the tongue, which are assumed to occur at a more superficial level, are not subject to lexical rules, but they are subject to postlexical processes; the forms created in word games, involving for instance modifications in the sequence of segments, undergo postlexical rules, but not lexical rules; lexical rules are not blocked by pauses, while postlexical processes usually are.

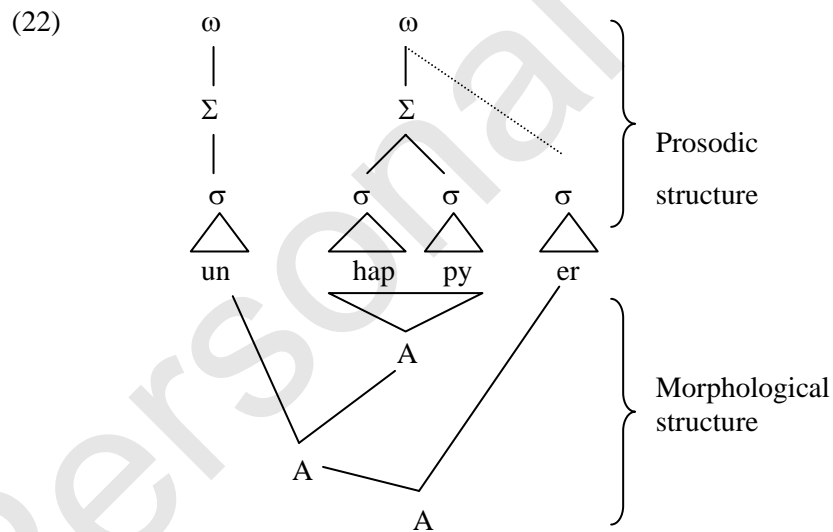
3.2. *The locus of construction of prosodic domains*

As mentioned above, the construction of prosodic domains was firstly seen to be accomplished postlexically (see paragraph 8). One problem of this approach, already anticipated in Nespor and Vogel (1986), results from the

fact that the construction of a prosodic domain such as the prosodic word may refer to morphological information, and this type of information is assumed no longer to be present at the input of the postlexical component. A solution to this type of problem is to allow for the construction of the lower prosodic domains in the lexical component. This has been proposed in Booij (1988), Booij and Lieber (1993), Inkelas (1990), Nespor (1990), and Vogel (1991), among others. Some of the arguments in favor of such a view are presented below.

There is evidence that morphological structure and prosodic structure must coexist (cf. Booij and Lieber 1993): for example, in Polish the choice between two particular allomorphs depends on whether the last consonant of the base word is syllabified by the syllabification algorithms of the language or it remains extrasyllabic; and reduplication in Tagalog requires affixation to a prosodic constituent in addition to a morphological constituent.

Another case where affixes have both morphological and prosodic subcategorizations is provided by the English comparative suffix *-er*. It subcategorizes for adjectival bases which, in addition, must be monosyllabic or disyllabic if the second syllable is light. The existence of forms such as *unhappier* suggest a paradox, since from a morphological point of view, the suffix is added to a trisyllabic base (*unhappy*).



16 *Theoretical background*

However, from a prosodic point of view the suffix attaches to a disyllabic prosodic word (*happy*), as the prefix *un-* constitutes an independent prosodic word (cf. Pesetsky 1985). These cases, therefore, do not constitute actual exceptions to the prosodic requirement if this requirement is imposed on the prosodic word rather than on the morphological base *-er* attaches to (cf. Booij and Rubach 1984). This implies the coexistence of morphological and prosodic planes. Under this approach a word such as *unhappier* has the structure in (22) (cf. Booij and Lieber 1993: 35).

Other elements that may have prosodic selection requirements are clitics (Inkelas 1990; Booij and Lieber 1993; Booij 1996a; Van der Leeuw 1997). For example, according to Booij and Lieber (1993) the Dutch clitic *ie* ('he') subcategorizes for a prosodic word preceding it. Notice, however, that this information is present in the lexical entry of the clitic, and therefore it does not imply that the prosodic structure is built in the lexical component. In fact, *ie* attaches to the prosodic word it selects only postlexically (cf. Berendsen 1986: chap.3; Booij 1996a).

Clitics provide, nevertheless, additional support for the presence of prosodic structure in the lexical level. This can also be illustrated by Dutch clitics, and their behavior with respect to final devoicing and resyllabification. According to Booij (1995, 1996a), among others, the prosodic word is the domain for resyllabification in Dutch. The fact that clitics such as *ie* are included in that domain shows that they are incorporated into the preceding prosodic word (cf. 23a). Syllable-final devoicing, a process whereby a voiced obstruent becomes voiceless in syllable final position (cf. 23b), is classified in Booij and Rubach (1987) as a word-level lexical rule, applying within a domain that includes lexical suffixes (cf. 23c). As illustrated in (23d), the postlexical integration of the enclitic into a prosodic word ending in an underlying voiced consonant, does not bleed final devoicing, although resyllabification causes the consonant to become syllable initial (see also Berendsen 1986, and Baumann 1996, who adds perceptual and acoustic evidence showing that the relevant resyllabified consonants surface unvoiced).

- | | | | |
|------|-------------------|------------|---------------------------------------|
| (23) | a. <i>komt-ie</i> | 'comes he' | (kɔm) _σ (ti) _σ |
| | b. <i>held</i> | 'hero' | (hɛlt) _σ |
| | c. <i>heldin</i> | 'heroine' | (hɛl) _σ (din) _σ |
| | d. <i>vond-ie</i> | 'found he' | (vɔn) _σ (ti) _σ |

The fact that lexical phonological processes refer to a prosodic domain, in this case the syllable, also confirms the existence of prosodic structure in the lexical component.

To sum up, the idea that prosodic domains up to the word level are built in the lexical component is supported at least by three major classes of arguments: (i) the observation that certain morphological operations refer to prosodic constituents; (ii) the fact that the construction of certain prosodic domains must refer to morphological information; (iii) the observation that some phonological processes refer to prosodic domains whose form may have to be defined prior to the concatenation of words and subsequent modifications (such as resyllabification and the integration of clitics).²³

4. The status of the Clitic Group within the prosodic hierarchy

The prosodic hierarchy is seen to include the Clitic Group since the proposal of Hayes (1989), adopted in Nespor and Vogel (1986) and much subsequent work (e.g. Vogel 1990, 1991; Hannahs 1995a; Nespor 1999a; Schwindt 2000, among others). This constituent groups together a prosodic word plus adjacent clitics, and is located above the prosodic word and below the phonological phrase. Like other prosodic domains, the clitic group has been reported to constitute the domain for many phonological rules. For example, according to Nespor and Vogel (1986) the clitic group is a domain for stress assignment in Latin. Thus, when enclitics are attached to a word, the primary stress is shifted from its original position in the word to the syllable that precedes the clitic, as illustrated in (24) (from Nespor and Vogel 1986: 146).

- (24) *vírum* *virúmque*
 ‘the man-acc’ ‘and the man-acc’
 vidēs *vidēsne*
 ‘you see’ ‘do you see?’

Other examples of phonological rules that have been argued to refer to the clitic group include *v*-Deletion and *s*, *z*-Palatalization in English (Hayes 1989), Demotic Greek Stress Readjustment, Italian Intervocalic *s*-Voicing, Greek Nasal Deletion, Nasal Assimilation and Stop Voicing, and Turkish Vowel Harmony (Nespor and Vogel 1986). The clitic group has also been

18 *Theoretical background*

proposed to be the domain for stress assignment within compounds in Hungarian (Vogel 1990).

There are, however, a number of important arguments against the existence of such a prosodic domain. We review some of these arguments in the next paragraphs.

According to Inkelas (1990), there seems to be no language where postlexical processes distinguish between the clitic group and the prosodic word; in most of the cases presented as evidence for the clitic group, the data can be reinterpreted as following from the distinction between lexical prosodic words and lexical rules versus postlexical prosodic words (that may include clitics) and postlexical rules. Thus, if the behavior of clitics can be explained by independently required tools, the clitic group should be eliminated on economy grounds (see also Booij 1988, 1996a; Selkirk 1996: note 3).

Under the proposals of Inkelas (1990), Inkelas and Zec (1991), Selkirk (1996), Kleinhenz (1996), Peperkamp (1997a), Hall (1999a), among others, clitics may attach not only to prosodic words, but also to prosodic phrases (see section 6.3). Thus, if it is assumed that prosodic hierarchy is universally defined, and the placement of the clitic group is fixed within that hierarchy, the distinction between word and phrasal clitics cannot be accounted for (cf. Peperkamp 1997a).

The definition of the clitic group presented in Nespor and Vogel (1986: 154) implies furthermore that clitics form independent prosodic words. Nevertheless, the most salient property of clitic words is their prosodic deficiency and the lack of properties characteristic of independent prosodic words. By abandoning the clitic group and the need for clitics to be independent prosodic words, it is possible to establish a one-to-one correspondence between the prosodic word and primary word stress: (i) each prosodic word must bear one main stress, and (ii) each stress bearing unit is contained within a prosodic word. Such statement could not be made under Nespor and Vogel's (1986) assumption that a prosodic word is defined as a constituent that bears at most one main stress.

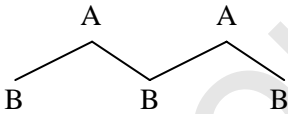
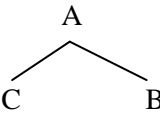
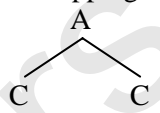
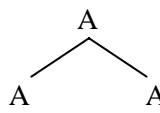
Finally, proclitics and enclitics (like prefixes and suffixes) often show asymmetries in terms of their coherence to the host (base) they attach to (e.g. Booij 1996a for Dutch, Peperkamp 1997a for Italian, Kleinhenz 1996 and Hall 1999a for German): enclitics (like suffixes) usually show a stronger degree of connection with their hosts (bases) when compared to proclitics (prefixes), which usually present a phonological behavior more independent of the host (base). Furthermore, enclitics and proclitics may also attach to different prosodic constituents (as in the case of German,

according to Kleinhenz 1996 and Hall 1999a). Under the clitic group approach, no such asymmetry should arise (cf. Kleinhenz 1996; Peperkamp 1997a).

In summary, there are strong reasons to assume that the clitic group is not a prosodic constituent. In section 6.3 we discuss the types of structural relations clitics may bear with their host, and we present some of the prosodic constituents that have been claimed to form possible hosts for clitics.

5. The Strict Layer Hypothesis

It has been observed that some essential properties distinguish prosodic structures from syntactic and metrical ones. The grouping of prosodic constituents is such that it forms n-ary structures, composed of a fixed number of constituents organized in a hierarchical fashion, with a limited depth. Consequently, the prosodic tree is flatter than syntactic and metrical trees. These properties are expressed by the Strict Layer Hypothesis (SLH), presented in (2) above (cf. Selkirk 1984; Nespor and Vogel 1986), which is assumed to constitute a set of well-formedness conditions on prosodic structure that define (im)possible geometric configurations of prosodic trees. The following structures, taken from Ladd (1996), illustrate some of the configurations that the SLH rules out (see Nespor and Vogel 1986: 8 for further exemplification on possible and impossible trees).

- (25) a. No multiple domination
- 
- b. No heterogeneous sisters
- 
- c. No skipping of levels
- 
- d. No recursion
- 

However, various phonological facts in the languages of the world have been reported to require analyses that violate certain aspects of the SLH. Some examples are briefly mentioned below.

Under Inkelas' (1990) theory, the satisfaction of the prosodic subcategorization frames of clitics always implies the violation of the SLH, since cli-

20 *Theoretical background*

tics are Chomsky-adjoined to the prosodic domain they subcategorize for. The analysis of prefixes and proclitics in Dutch and in several dialects of Italian, developed in Booij (1996a) and Peperkamp (1997a), respectively, also implies the violation of the SLH in that these elements may form syllables which are adjoined at the level of the prosodic word or the phonological phrase. Finally, Leben and Ahoua (1997) suggest that a number of structures in Baule – a language spoken in Côte d'Ivoire – form prosodic words that may have internal prosodic words. Thus, recursion is implicitly admitted at the level of the prosodic word (see also section 6.4).

Recursion at the level of the intonational phrase level has also been argued for, namely by Ladd (1992, 1996) and Frota (2000).

Ladd (1992) questions the Strict Layer Hypothesis on the basis of arguments of various sorts: (i) the more closely phonological phenomena are studied, the greater tendency there is to introduce new prosodic domains – which suggests a prosodic structure with no fixed depth; (ii) the fixed depth of prosodic structure is not always supported by phonetic cues of boundary strength; (iii) certain prosodic groupings of intonational phrases that disambiguate sentences may violate the SLH – for example, a sentence like the one in (26) may be realized so that [*his faithful black labrador*] is interpreted as one of the elements Dubois lives with (in which case it constitutes an intonational phrase of the same type as adjacent intonational phrases), or it may be realized so that it is interpreted as the same entity as Jean-Charles (in which case there is a stronger prosodic connection between this intonational phrase and the preceding one).

- (26) *Dubois lives in a restored 15th century farmhouse with Jean-Charles, his faithful black Labrador, and a motley assortment of cats.*

In order to account for such facts, Ladd puts forth the Compound Domain Hypothesis, according to which prosodic domains – at least at the intonational phrase level – are allowed to be grouped together under a prosodic constituent of the same level.

Frota (2000) has also shown that a number of phonological phenomena require (limited) recursion at the intonational phrase level in EP.²⁴ In particular, Frota shows that several phonological phenomena point to the existence of an intonational phrase domain (I^{\max}) that dominates two constituents of the same type (which we will refer to as I^{\min}). These phenomena include (i) sandhi processes – for example, fricative voicing is bound by the I^{\max} -domain, whereas consonant deletion involved in the process of syllable degemination is bound by the I^{\min} -domain; (ii)

preboundary lengthening – despite the existence of significant preboundary lengthening at the right edge of both the I^{\min} -domain and the I^{\max} -domain, the amount of preboundary lengthening in the former is significantly shorter than the one found in the latter; (iii) pause distribution – acoustic pauses seem to signal I^{\max} boundaries, rather than I^{\min} boundaries; (iv) F0 range – although both I^{\max} and I^{\min} are tonally marked with the nuclear tone and a boundary tone (which are identifying features of the I-domain in EP), the tonal events associated with an I^{\max} are realized with a wider pitch range, than those associated with an I^{\min} .²⁵

A number of proposals have therefore been put forward in order to weaken or relax the requirements embodied by the SLH (e.g. Ladd 1992; Selkirk 1996). Along the lines of Inkelas (1990) and Itô and Mester (1992), Selkirk (1996) proposes that the SLH should be decomposed into a set of constraints on prosodic domination, formulated and exemplified in (27) (where C^n stands for some prosodic category).

- (27) a. *Layeredness*
 No C^i dominates a C^j , $j > i$ (e.g. no σ dominates a Σ)
 b. *Headedness*
 Any C^i must dominate a C^{i-1} , except if $C^i = \sigma$ (e.g. a PW must dominate a Σ)
 c. *Exhaustivity*
 No C^i immediately dominates a constituent C^j , $j < i-1$ (e.g. no PW immediately dominates a σ)
 d. *Nonrecursivity*
 No C^i dominates C^j , $j = i$ (e.g. No Σ dominates Σ)

Among these, the first two are argued to hold universally and to be unviolable, whereas Exhaustivity and Nonrecursivity are considered to be violable constraints (as the exemplification in the preceding paragraphs has already shown).²⁶

Although weakened, the modifications introduced in the SLH are such that it still captures the original idea that prosodic structures are flatter than syntactic structures, and are composed of a fixed prosodic constituency.²⁷

22 *Theoretical background***6. The prosodization of words***6.1. Diagnostics for the prosodic word*

Primary word stress is one of the most intuitive diagnostics for the prosodic word domain. In fact, it is generally accepted that the prosodic word must bear one and only one primary stress. Adopting the view that clitics cannot form independent prosodic words, the generalization in (28) can be formulated (see also Van der Leeuw 1997, among others).²⁸

(28) A prosodic word must bear one and only one primary (word) stress

Like other domains, the prosodic word is a domain for the *application of phonological rules*. For example, Nespor and Vogel (1986) present several processes of various languages that are bound by the prosodic word domain: namely, Nasal Assimilation and Stop Voicing in Greek, Main Stress Rule in Latin, Final Voicing in Sanskrit, and Vowel Harmony in Turkish.

In addition to word stress and (other) phonological rules, the prosodic word has also been reported to be a domain for *phonotactic generalizations*. To cite just a few examples, in Dutch, the right edge of the prosodic word allows for syllables longer than those found word internally (Booij 1995), in Italian, prosodic words cannot begin with [ʎ] (Peperkamp 1997a), in German, short full lax non-low vowels cannot occur at the right edge of the prosodic word (Hall 1999a), in English, there are more consonantal clusters word internally than at either edge of the prosodic word domain (Raffelsiefen 1999a).²⁹

In pitch accent languages such as Serbo-Croatian, pitch accents are assigned to every prosodic word but not to clitics (e.g. Godjevac 2000). In such languages, therefore, this type of tonal information may be used as diagnostics to the prosodic word as well.³⁰

Deletion under identity may also cue the prosodic word. In fact, Booij (1985, 1988), Wiese (1993, 1996), Kleinhenz (1994) propose that in Dutch and German the deletion of an element within complex words in partially similar coordinate structures depends not only on morphosyntactic information, but also on the prosodic status of the element to be omitted in the string: besides partial phonological identity with respect to the other element of the coordinate structure, it must also form an independent prosodic word.³¹ Representative examples from Dutch that show the relevance for coordination reduction of the phonological word rather than the morpho-

syntactic word are presented in (29) (taken from Booij 1985).³² In (29a), both *schei* and *kunde* constitute prosodic words, but the former element does not exist as an independent word. In (29b) the suffix *-achtig* may be deleted because it forms a prosodic word, contrasting with *-ig*, in (29b'), which does not form an independent prosodic word and therefore may not undergo coordination reduction.³³

- (29) a. ~~scheikunde~~ en natuurkunde > *schei en natuurkunde*
 analysis knowledge and nature knowledge
 'chemistry and physics'
 b. storm~~achtig~~ en regenachtig > *storm en regenachtig*
 'stormy and rainy'
 b'. blauw~~ig~~ en rodig > **blauw en rodig*
 'bluish and reddish'

A wide number of languages have additionally been reported to show *Minimal Word requirements*, that is, the prosodic word is argued to have a minimal size, usually being at least disyllabic or bimoraic. According to the literature surveyed in Kenstowicz (1994), such languages include English, Yidin^y, Arabic, Japanese, Lardil, Estonian, and Choctow. To this list we can add Bengali (Fitzpatrick-Cole 1991), Catalan (Cabr  1993), Baule (Leben and Ahoua 1997), German (Hall 1999a), many Bantu languages cited in Downing (1999) and Chamicuro (an Amazonian language – cf. Parker 1999), among others. Notice, nevertheless, that a few languages do not seem to show the minimal word syndrome, namely, Irish (cf. Green 1997, cited in Hall 1999b), and Brazilian Portuguese (cf. Bisol 2000). In the latter language (as well as in European Portuguese), a prosodic word may consist of a single syllable, whether closed or open, and headed by low, mid or high vowels. This is illustrated in (30), taken from Bisol (2000).³⁴

- (30) *tr s* *p * *mi*
 'three' 'foot' 'mi'
 pai *p * *nu*
 'father' 'dust' 'naked'

Clipping (or truncation) is another process that may provide evidence for the prosodic word domain. In fact, various languages have been reported to have morphological operations consisting of the shortening of words whose output forms a (minimal) prosodic word (e.g. Mester 1990 for

24 *Theoretical background*

Japanese; Prieto 1992 for Spanish; Cabré 1993 and Cabré and Kenstowicz 1995 for Catalan; Thornton 1996 for Italian). Examples of truncation in Italian are provided in (31) (from Thornton 1996: 87).

- (31) a. *amplifictore* (*ampli+fica+tore*) > *ampli* 'amplifier'
 b. *diapositive* (*dia+positiva*) > *diapo* 'slide'
 c. *meteorologico* (*meteo+logico*) > *meteo* 'concerning weather'

Notice that in some of these cases the clipped form (which retains the semantics of the original base) does not correspond to a morphological constituent (cf. 31b–c). By contrast, in all cases it corresponds to a minimal prosodic word, which Thornton argues to be the disyllabic trochee in Italian.

In Germanic languages, where (re)syllabification is bound by a word-size constituent, the domain of *syllabification* is also usually taken to be a reliable diagnostic for prosodic wordhood (e.g. Booij 1995, 1996a, for Dutch; Wiese 1996, Hall 1999a, for German; Raffelsiefen 1999a, for English). However, in Romance languages, the picture is not so clear due to the existence of syllabification at the word level and of resyllabification across words (see, for example, the discussion in Nespor and Vogel 1986: 3.1.1, and section 7 below).³⁵

To conclude, several types of phenomena may cue the prosodic word domain. Nevertheless, it is possible that only a subset of these diagnostics may play a role in the identification of prosodic words in each language. This is, for instance, the case of EP, as we will show in chapter 5.

6.2. *The prosodic word domain*

According to Nespor and Vogel (1986: chap.4), the prosodic word (PW) may be either of the same size or smaller than a syntactic terminal node, depending on language-particular definitions of the prosodic word domain. In languages such as Greek and Latin there is a coincidence between the prosodic word domain and the terminal syntactic node: a PW includes a stem, plus all adjacent affixes, as well as both members of compounds. In languages such as Turkish or Sanskrit, by contrast the PW domain is smaller than a syntactic terminal node, since each member of a compound structure functions as a prosodic word domain.

In other languages, additional information may be required in order to define the PW domain. This is the case of Hungarian, where prefixes seem to form their own prosodic word, as well as of Italian, where only certain prefixes, depending on phonological properties (namely, syllable structure), form independent prosodic words, and Dutch, where certain suffixes form their own prosodic word and have to be lexically marked as prosodic words.³⁶

In order to account for the different types of languages observed, Nespor and Vogel (1986) propose the definitions of the prosodic word domain in (32).³⁷

(32) PW domain (two major types of languages)

Type I. The domain of PW is Q (Q=syntactic terminal node)
(e.g. Latin)

Type II. The domain of PW consists of (a) a stem; (b) any element identified by specific phonological and/or morphological criteria (e.g. Italian); (c) any element marked with the diacritic [+W] (e.g. Dutch). Any unattached elements within Q form part of the adjacent PW closest to the stem.

More recently, some aspects of Nespor and Vogel's proposal have been either questioned or refined. For example, unlike Nespor and Vogel, and adopting the view that clitics do not constitute independent prosodic words, Booij (1996a), among others, shows that the combination of a host and a clitic constitutes a case where a prosodic word may correspond to a unit larger than a syntactic terminal node. In fact, since phonological clitics may behave syntactically in the same way as phonologically non-clitic words, it may be assumed that (at least certain) clitics constitute syntactic terminal nodes distinct from the terminal node corresponding to their host (the issues related to the prosodization of clitics are treated in the following section).

Furthermore, it is shown in Nespor and Ralli (1996) that compound words both in Greek and Italian may be computed either as a single prosodic word, or as two prosodic words. According to these authors, the choice between the two possibilities depends on the morphological structure of the compound. This suggests that the distinction between the two classes of languages referred above is (at least) not supported by Greek and Italian data (the prosodization of compounds is discussed in more detail in section 6.4, below).

26 *Theoretical background*

Finally, Peperkamp (1997a) proposes an alternative analysis of Italian prefixed words, in which prefixes are claimed not to incorporate into the base prosodic word, regardless of syllabic structure considerations (the prosodization of affixes is briefly considered in section 6.5 and the issues related to the (re)syllabification of prefixes are discussed in section 7).

To conclude, although it is clear that there must be a word-sized prosodic domain distinct from any morphosyntactic constituent, the definition of the prosodic word in the languages of the world is still an open matter. The following sections are devoted to a more detailed inspection of some of the questions raised by the prosodization of clitics, compounds and derived words.

6.3. *The prosodization of clitics*

Of all topics discussed so far, clitics and related matters are among those that have attracted the greatest attention from researchers (e.g. Zwicky 1977; Klavans 1985; Neijt 1985; Berendsen 1986; Inkelas 1990; Zec and Inkelas 1991; Selkirk 1996; Booij 1996a; Kleinhenz 1996; Peperkamp 1996, 1997a; Van der Leeuw 1997; Nespor 1999a; Hall 1999a; Bisol 2000, among many others). In this section we will concentrate on some of the core questions related to the prosodization of clitics.

Assuming that the clitic group is not a prosodic constituent, phonological clitics must attach to items that correspond to some other prosodic domain. Thus, the first question that we have to address concerns the identification of the possible *prosodic hosts* for clitics.

In the literature where the clitic group is excluded from the prosodic hierarchy, we find several possible prosodic hosts, ranging from the prosodic word, to the phonological phrase, and the intonational phrase. For example, in the work of Zec and Inkelas, the three constituents are argued to constitute possible hosts for clitics in different languages: clitics attach to the prosodic word both in Modern Greek and in Serbo-Croatian (Zec and Inkelas 1991); the emphatic particle in Hausa, as well as determiners in Kivunjo Chaga (a Bantu language), attach to the ϕ -phrase (Inkelas 1990); clitics attach to the I node in languages like Tzotzil (a Mayan language) (Zec and Inkelas 1991). We will mention just two types of evidence for these proposals. In Kivunjo Chaga there are phrasal phonological rules that show the presence of a ϕ -boundary between the phonological host (a noun) and the clitic (a determiner). In the case of Hausa and Tzotzil the argument is based on the distribution of clitics, rather than on direct phonological evidence. It

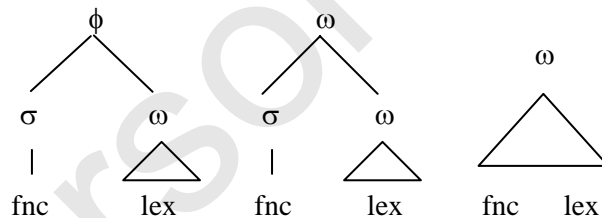
is claimed that the relevant clitics are located with reference to prosodic constituents rather than to syntactic constituents, and a direct correlation is established between the prosodic constituent with respect to which the clitic is located and its prosodic host (see, nevertheless, Van der Leeuw 1997: 2.6 for a critical discussion of some of these proposals).

Other examples of phrasal hosts for clitics are found, for example, in the analysis of English clitic function words in phrasal non-final position proposed in Selkirk (1996), in the analysis of Standard Italian pronominal clitics in Peperkamp (1997a), or the analysis of German proclitics in Kleinhenz (1996) and Hall (1999a). It appears, nonetheless, that the hosts of clitics correspond more commonly to prosodic words. This is the case of English reduced object pronouns (Selkirk 1996), of Dutch clitics (Booij 1996a), of pronominal clitics in Neapolitan and Lucanian (Peperkamp 1997a), as well as of most German enclitics (Kleinhenz 1996; Hall 1999a), and of Chamicuro determiners (Parker 1999).

Not only can languages and dialects vary according to the possible hosts for clitics, but they can also vary in the way clitics are integrated in prosodic structure (e.g. Selkirk 1996; Booij 1996a; Peperkamp 1997a; Hall 1999a).

According to Selkirk (1996), who assumes that Exhaustivity and Nonrecursivity are violable constraints (as we have seen in section 5), clitic function words may be prosodized in three different ways, as represented in (33) (where *fnc* and *lex* correspond to the phonological content of function words and lexical words). The different types of configurations where clitics may appear with respect to the prosodic word give rise to their classification as “free clitics”, “affixal clitics”, and “internal clitics”.

(33) a. free clitics b. affixal clitics c. internal clitics



An example of the instantiation of each possibility is provided by three NeoŠtokavian dialects of Serbo-Croatian (cf. Zec 1993, reviewed in Selkirk 1996). In each dialect, a default initial high tone accent is assigned to the first mora of the prosodic word. The realization of the initial accent varies,

28 *Theoretical background*

however, depending on the way the clitic is structured with respect to the host: in the dialect of E. Herzegovina (NŠ-1), the initial accent is realized on the first mora of the clitic, and not on the host lexical word; in the (standard) dialect of Belgrade (NŠ-2), the initial accent is realized on the first mora of the lexical word, but not of the clitic; in the dialect of Šrem, Mačva (NŠ-3) there are two possibilities – the high accent may fall on the first mora of the clitic, or it may fall on the first mora of the lexical word, in which case it spreads to the preceding clitic. Thus, in NŠ-1 the clitic behaves as PW initial, as in the representation in (34a); in NŠ-2 the clitic behaves as PW external, as in the representation in (34b); and finally, in NŠ-3 both the clitic and the lexical word may receive the initial accent, a possibility that would follow from a representation like (34c) (in the examples the acute accents represent the high tone accents).

- (34) a. internal clitic (ú graad)_ω ‘to the city’
 b. free clitic u (gráad)_ω ‘to the city’
 c. affixal clitic (ú (graavu)_ω)_ω ‘into (the) head’
 or (ú (gráavu)_ω)_ω

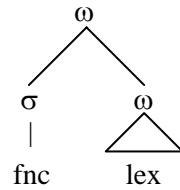
On the basis of word stress assignment, Peperkamp (1997a) argues for a similar kind of distinction between different Italian dialects (or languages). Observing that in certain dialects of Lucanian enclitics induce the main stress to shift to the penultimate syllable, the author proposes that in these dialects enclitics are incorporated into the host prosodic word, as in (35a). In Neapolitan, by contrast, both the host and the first of two pronominal enclitics are assigned main stress, which is accounted for by assuming that the sequence host plus clitic is prosodized in a recursive structure, as represented in (35b).³⁸ Finally, in Standard Italian enclitics do not interact with main stress assignment and this is taken to indicate that they are prosodized in a position external to the prosodic word, as in (35c) (in the examples acute stress represents main word stress).³⁹

- (35) a. Lucanian (man:atə míllə)_ω ‘send it to me’
 b. Neapolitan ((cóna)_ω tíllə)_ω ‘tell it to you’
 c. Standard Italian ((pórta)_ω melo)_φ ‘bring it to me’

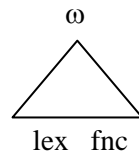
The variation in the prosodization of clitics may also be found within the same language/dialect, depending either on the direction of cliticization

(cf. Booij 1996a) or on the segmental string that results from the combination of lexical and clitic words (cf. Hall 1999a).⁴⁰ Specifically, based on different phonological phenomena that apply with reference to the prosodic word domain – namely, resyllabification, prevocalic schwa deletion, /n/-insertion, and homorganic glide-insertion – Booij (1996a) shows that Dutch *proclitics are adjoined* to the following prosodic word, as in the representation in (36a), whereas *enclitics are incorporated* in the preceding prosodic word, as in the representation in (36b).

(36) a. Dutch proclitics



b. Dutch enclitics

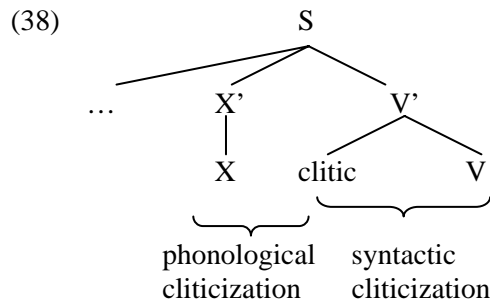


Hall (1999a), in turn, proposes that enclitics in German may either be incorporated into the preceding prosodic word (as in Dutch), or be prosodized in a position external to the prosodic word if the result of the combination of the clitic with the preceding word violates the language phonotactic constraints on the prosodic word. For example, consonant-initial enclitics may not be incorporated into the preceding prosodic word when the enclitic ends in a short full lax vowel, because this would violate the Lax Vowel Constraint that prohibits prosodic words to end with such a vowel. Hall thus proposes the two possible representations in (37) for German enclitics.⁴¹

(37) German enclitics (host+enclitic)_ω or ((host)_ω + enclitic)_φ

A final major issue related to phonological cliticization concerns the *direction of cliticization*.

Ever since the influential work of Klavans (1985), it has been well established that there is no necessary correlation between the direction of syntactic attachment of clitics, and the direction of phonological cliticization. An example where the direction of the two types of attachment does not coincide is provided by Nganhacara (an Australian language), where clitics attach phonologically to their left, while their syntactic host may appear to their right. This is schematized in (38) (adapted from Klavans 1985: 105).

30 *Theoretical background*

The question then arises of how the directionality of cliticization is defined, since it is not sufficient to assume, like in Hayes (1989: 208) for English, that a clitic attaches to the constituent with which it shares more category membership (where ‘X and Y share category membership in C if C dominates both X and Y’).

Along the lines of Klavans (1985), Nespor and Vogel (1986) claim that syntactic configuration is not enough to determine the direction of phonological cliticization, and that it may follow, instead, from an inherent property of individual clitics. The relevance of the inherent property of individual clitics is demonstrated by the Greek possessive clitics, which, like Nghanacara clitics, are phonological enclitics, although syntactically proclitic.

A slightly different approach is proposed in Booij (1996a). For this author, languages may show a preference for a given direction of cliticization, which may follow from the language’s rule of Stray Adjunction (Anderson 1992: 203). Individual clitics may, in addition, be specified to cliticize only to the right or to the left. Thus, in the case of Dutch the preferred direction of cliticization is leftwards, although rightward cliticization is also possible, for example, if no host occurs to the left of the clitic. The clitic *-ie*, however, is specified to cliticize phonologically only to the left.⁴² In the same line of approach, Kleinhenz (1996) suggests that German displays a similar preference for leftward cliticization, and the possibility of bi-directional cliticization, maintaining, further, that in this language no clitics appear to attach solely to the right.

6.4. The prosodization of compounds

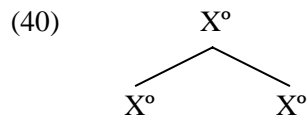
The prosodization of compound structures has been recently studied in depth in Nespor and Ralli (1996), Peperkamp (1997a, 1997b), and Nespor (1999b).

Nespor and Ralli (1996) and Nespor (1999b) establish the following correlation between the morphosyntactic structure of Greek compounds and primary stress assignment: (i) in *stem plus stem* compounds, a single stress falls on the antepenultimate syllable of the compound; (ii) in *stem plus word* compounds, there is also a single stress, but in this case it is located in the same place as when the relevant word occurs in isolation; (iii) in *word plus word* compounds, there are two main stresses which fall in the same place as when each of the words occurs in isolation. Examples of each type are given in (39) (from Nespor and Ralli 1996).

(39) a. stem+stem	<i>tirópita</i>	<	tir–	pit–
	‘cheese pie’		‘cheese’	‘pie’
b. stem+word	<i>tamosaláta</i>	<	tam–	saláta
	‘caviar salad’		‘caviar’	‘salad’
c. word+word	<i>nómos plésio</i>	<	nómos	plésio
	‘law-frame’		‘law’	‘frame’

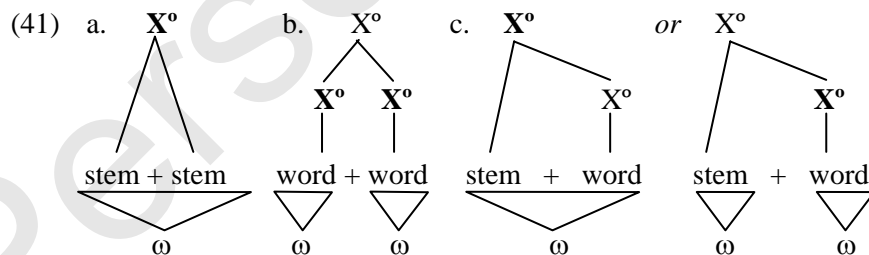
These facts lead the authors to propose that the first two types of compounds are prosodized as a single prosodic word, whereas the third type corresponds to two prosodic words that, together, form a phonological phrase. Besides word-stress, the analysis is corroborated, in particular, by two additional phonological rules of Greek – progressive nasal assimilation, and regressive stop voicing assimilation – which do not apply across prosodic words. The authors suggest, further, that the variation in the way compounds are prosodized – as a prosodic word, or as a phonological phrase – follows from the place in the grammar where these constructions are formed. Stem+stem and stem+word compounds are assumed to be formed in the lexical level, while word+word compounds are generated as syntactic phrases and are reanalyzed as X^0 categories when they reenter the lexicon (cf. Di Sciullo and Williams 1987).

According to Vogel (1990), the way compounds are prosodized may also vary depending on the way languages construct the clitic group domain. Given the general structure of word compounds in (40), a language may select either the lowest or the highest X^0 for the formation of this domain.

32 *Theoretical background*

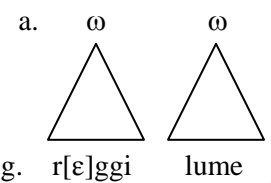
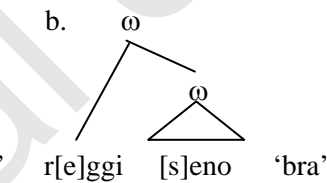
In contrast with this view, Nespor and Ralli (1996) and Nespor (1999b) propose that the prosodic domains that are relevant for the prosodization of compounds are the prosodic word and the phonological phrase. In addition, they show that only the lowest X^0 seems to be selected in the mapping of compounds onto prosodic words. Thus, in the case of stem+stem compounds, the lowest X^0 corresponds to the node that dominates the two stems (since stems are not X^0 level elements), and it is therefore predicted that this type of compound always forms a single prosodic word.⁴³ In the case of word+word compounds, the lowest X^0 corresponds to each member of the compound, and therefore each element is mapped onto an independent prosodic word.⁴⁴ Finally, in the case of stem+word compounds, some variation seems to be allowed, depending on the relevance assigned to either the left or the right member of the compound: in Greek, in some varieties of Italian, and in Icelandic, the whole compound is mapped onto a single prosodic word, thus suggesting that the lowest X^0 is determined with reference to the first constituent; in other Italian varieties, as well as in Dutch, this type of compound is mapped onto two prosodic words, thus suggesting that the lowest X^0 is determined with reference to the second member of the compound.⁴⁵ As for the prosodization of the first member of the compound as a prosodic word in the latter case, it is proposed that it follows from the requirements of the Strict Layer Hypothesis (which is assumed to be unviolable).

The different possibilities for the prosodization of compounds are represented in (41) (adapted from Nespor 1999b: 139), where the X^0 that counts for the construction of the prosodic word domain is signaled in boldface.



As said above, both in Nespor and Ralli (1996) and in Nespor (1999b) it is assumed that word+word compounds are grouped together in a phonological phrase. Evidence for this comes from the relative prominence of the primary stresses assigned to each prosodic word. Whereas languages may vary according to the element of the compound that is stronger, there is always one element that bears a stronger stress, and this element is taken to be the head of the phonological phrase.

Within an optimality theoretical approach (cf. Prince and Smolensky 1993), and assuming a relaxed version of the Strict Layer Hypothesis, Peperkamp (1997a) claims that Italian word+word compounds may be prosodized as in (42a). If the word is frequent in the lexicon of individual speakers, however, it may be structured in a recursive structure of the type shown in (42b).⁴⁶ Evidence for the latter structure is provided by two Italian phenomena, illustrated by the example in (42b). That the first member of the compound lacks prosodic word status is shown by the fact that the vowel that corresponds to the stressed element when the word occurs in isolation undergoes vowel raising, a process that applies to unstressed vowels. The fact that there is no intervocalic *s*-voicing, a process that is assumed to apply in prosodic word internal position (Nespor and Vogel 1986), shows that the fricative is still treated as prosodic word initial.

- (42) a.  b. 

It should be noticed that Peperkamp (1997a) adopts a restrictive view of the notion of compounding, in that the only productive word+word combination in Italian is considered to be of the form verb+noun, whereas other types of combinations are assumed to be constructed as syntactic phrases which have undergone a lexicalization process.⁴⁷


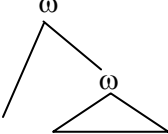

Unlike Nespor and Ralli (1996) and Nespor (1999b), Peperkamp does not take a position concerning the prosodic constituent that dominates compounds that include more than a single prosodic word. By contrast, Leben and Ahoua (1997) suggest that in Baule several word+word compounds that behave as grouped within a single prosodic word may contain internal prosodic words. It is therefore implicit in this work that compounds may display a recursive structure. Hannahs (1995a: 3.4.2) also shows that

34 *Theoretical background*

word compounds in French pattern differently from non-compound words included within a phonological phrase. This leads him to propose that word compounds in French are grouped within a clitic group, along the lines of Vogel (1990).

6.5. *The prosodization of derived words*

As we have already seen in section 6.2, there seems to exist both cross-linguistic and intra-linguistic variation in the way (some) prefixes and (some) suffixes are prosodized. In fact, a survey of the literature on the issue shows that derived words may behave prosodically like simplex words, like certain host-clitic sequences, or like adjacent prosodic words. Examples of prefixed words in Italian of each type are presented in (43).

- (43) a. 
resistenza
'resistance'
- b. 
pre finanziare
'to prefinance'
- c. 
super attivo
'superactive'

(Nespor and Vogel 1986:127) (Peperkamp 1997a:3.3.1) (Peperkamp 1997a:3.3.1)

Among the sources of variation in the way affixes are prosodized are the phonological make-up of certain affixes. For example, in some languages suffixes may form independent prosodic words depending on their phonological properties. This seems to be the case of Dutch *-achtig* '–like', *-baar* '–able', *-heid* '–ness', and four other suffixes, listed in Booij (1995: 111). In fact, according to this author, the suffixes that form independent prosodic words are almost totally predictable, since they consist of the native suffixes that have the phonological properties of phonological words.⁴⁸ In German, Wiese (1996: 3.4) proposes that all morphemes – including suffixes – starting in a consonant and followed by a vowel form an independent prosodic word (see also Raffelsiefen 1999b). Notice that, in this respect dialectal variation seems also to be allowed, since it is observed that in the Southern varieties of German consonant-initial suffixes do not constitute independent prosodic words, according to the syllabification test for prosodic wordhood.⁴⁹

The prosodic status of prefixes is also often seen to follow from the language Minimal Word requirement. This is proposed in Peperkamp (1997a) for Italian and Spanish, where disyllabic prefixes, unlike monosyllabic prefixes, are assumed to form independent prosodic words. The same type of requirement is shown to play a crucial role in the prosodization of prefixes in German (cf. Hall 1999a), and in English (cf. Raffelsiefen 1999a).

An important additional factor that has been argued to be involved in the way affixes – in particular prefixes – are prosodized has to do with their *productivity* and/or the *transparency* of the complex morphological structure. In fact, non-productive/non-transparent prefixes are often reported to be incorporated into the prosodic word that includes the morphological base (e.g. Nespor 1985, Nespor and Vogel 1986: 127, Peperkamp 1997a, and the example in 43a, above, for Italian; Booij 1995: 120, for Dutch; Hannahs 1995a, for French; Raffelsiefen 1999a, for English).⁵⁰

7. Syllabification, resyllabification, and prosodic word restructuring

Phrasal resyllabification seems to be a property common to Romance languages (e.g. Italian, Spanish, French, Catalan, Galician, EP), but not Germanic languages (e.g. English, Dutch, German).⁵¹ In fact, in Romance languages a prosodic word final consonant may appear at the onset of the following vowel-initial prosodic word, whereas this does not occur in Germanic languages. The example (44) illustrates resyllabification in Catalan (adapted from Lloret 2002: 242–243).

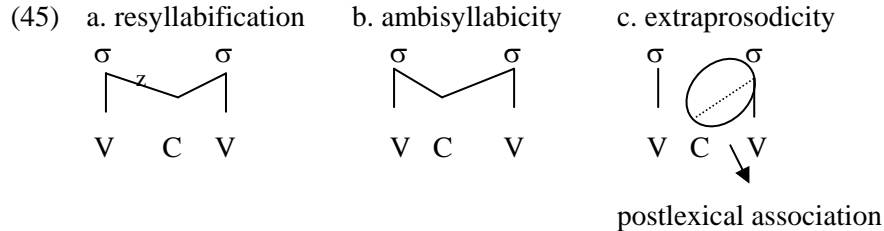
(44)	Syllabification of words in isolation	Resyllabification
	(son) _σ ‘are’	(so) _σ (ne) _σ (lles) _σ
	(e) _σ (lles) _σ ‘them’	‘it is them’

As the discussion below shows, the relevant distinction seems to be in the domain of resyllabification, rather than in the absence of resyllabification in certain languages.

Based on economy principles, linguists have tried to eliminate resyllabification as a structure changing rule, as in the representation in (45a). With this goal, Rice (1989) proposes to account for some of the cases previously considered to involve resyllabification as deriving from either ambisyllabicity or from extraprosodicity.⁵² In the former case, final consonants are suggested to be linked both to the preceding coda and to the following onset, as represented in (45b). In the latter case, word final conso-

36 *Theoretical background*

nants are assumed to remain lexically unparsed – that is, not to be attached to the syllable node at the lexical level – and to be syllabified only postlexically, as represented in (45c).⁵³



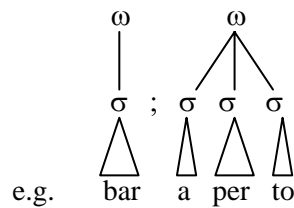
Extraprosodicity has been challenged with data from various languages on the basis of processes that apply before resyllabification, and that refer to segments that are assumed to be extraprosodic (cf. Booij 1994; Hall 1994; Peperkamp 1997a). Booij (1994) and Peperkamp (1997a: 2.2.3.1) further show that in languages like Dutch and Italian, respectively, there is evidence that ambisyllabicity is not a possible analysis for resyllabification.

An example from Spanish illustrates the inadequacy of the analysis of resyllabification resorting to the extraprosodicity of final consonants (cf. Hall 1994). In this language, the rules of Velarization and Aspiration apply to /n/ and /s/, respectively, when these consonants appear in syllable final position (cf. Harris 1983). The relevant segments surface as velar/aspirated, even if resyllabification has caused these consonants to surface in the onset of the following syllable. Thus, in order for these syllable final processes to apply, the relevant consonants must have been associated to a coda position, before postlexical resyllabification has applied (see also Nespor and Vogel 1986: 69–70).

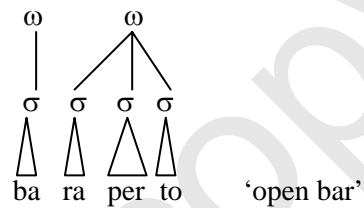
A case where the ambisyllabicity approach is shown not to be a viable analysis for resyllabification effects is provided by Dutch (Booij 1994). As already mentioned in section 3.2, syllable final obstruents in Dutch are devoiced. The postlexical attachment of vowel initial clitics induces to resyllabification, since clitics are incorporated into the preceding prosodic word, and the prosodic word is the domain of resyllabification in this language (see also section 3.2). As in the Spanish case, even if the obstruent occurs at the onset of the following syllable, it surfaces unvoiced. The crucial point here is that it is not possible to explain final devoicing by postulating that the relevant consonant is ambisyllabic, since in word internal position ambisyllabic obstruents do not trigger devoicing (cf. Booij 1995: 3.4).

Resyllabification as a structure changing process seems therefore to be required in order to account for the effects reported above.⁵⁴ This means that a consonant that ends a prosodic word at the output of the lexical component may be resyllabified, so that it becomes part of the next prosodic word postlexically. In order to avoid ill-formed prosodic representations induced by resyllabification (e.g. involving multiple domination), Peperkamp (1997a) proposes that the prosodic word boundaries lexically obtained, like in the example in (46a), are postlexically readjusted when resyllabification applies, yielding a structure like the one in (46b).

(46) a. Lexical prosodization



b. Postlexical prosodic word boundaries readjustment



Since morphological structure and prosodic structure belong to separate plains, the resulting mismatch in the alignment of word boundaries is assumed not to induce an ill-formed representation.⁵⁵

8. Concluding remarks and outline of the book

We may conclude from this presentation that a number of issues related in particular to the possible ways for morphemes and words to be integrated into the prosodic structure are still open issues. The aim of the present book is to contribute to the understanding of some of these issues, on the basis of European Portuguese data.

In section 6 it was seen that not only languages, but also dialects, may vary in the way affixes and clitics are prosodized. The *data* considered may thus be of crucial importance in this domain of research. In order to control for dialectal variation, namely the (yet unstudied) possibility that the same kind of elements may be prosodized differently in distinct dialects of EP, and thus, yield contradictory phonological behavior, we have restricted the scope of our investigation exclusively to a single dialect.⁵⁶ Specifically, we describe throughout this work the standard variety of EP spoken in the region of Lisbon, by educated speakers from 20 to 40 years. Although the

38 *Theoretical background*

data considered come in many cases from introspection, we have (informally) consulted other speakers of the same variety of EP in those cases where judgments may be less clear. In addition, we have collected external data on some of the processes considered in chapter 3. This was generally done in an informal way, either through the direct inspection of other speakers' speech, in conversation or through the observation of productions heard in the media (television or radio). There was, nevertheless, a number of facts, in particular those related to the reduction of function words, for which data were collected in a formal experimental way. These phenomena were, in fact, particularly difficult to describe merely on the basis of either intuition or informal observation. The procedure followed for collecting the data, as well as the identification of the relevant phenomena and their description and analysis is reported in chapter 7.

In section 3 we saw that phonological phenomena may be active only within the lexical component, and thus they may not be affected by any contextual modifications triggered by the postlexical combination of words. In order to assess how the elements that are postlexically combined are prosodized, it is therefore crucial to use as diagnostics for prosodization only processes that are operative postlexically.⁵⁷ Since the assessment of the prosodization of such postlexically combined elements is one of the major goals of the present work, it is essential for us to determine which processes may constitute reliable diagnostics for the postlexical prosodization of words, as well as for the prosodic word domain. In order to achieve this goal, chapter 3 will be devoted to the description and classification of the processes assumed in subsequent chapters as either lexical or postlexical, and as referring or not to the prosodic word domain in their structural description.

The relevance of the investigation conducted in chapter 3 for the present investigation is in fact twofold. On the one hand, the processes classified as postlexical are those which will be used in chapters 5 and 6 as diagnostics for the prosodization of elements combined postlexically. On the other hand, the processes classified as operating within the lexical component play an important role for determining whether pronominal clitics in EP behave like affixes combined with their base in the lexical component (like inflexion), or like clitics combined with their hosts postlexically (like clitic function words). The latter issue, namely the affix/clitic status of pronominal clitics, is the topic of investigation in chapter 4.

The core chapters of this book are chapters 5 and 6. Here we investigate how different kinds of morphosyntactic entities (affixes, clitics, and words) are integrated into prosodic structure at the level of the prosodic word.

As we have seen in section 6.1, the diagnostics for the prosodic word seem to vary crosslinguistically. Crosslinguistic variation may result from a difference in the domain of a given phenomenon (as in the case of resyllabification in Germanic versus Romance languages, reported in section 7); it may also result from the (non-)operation of a given phenomenon postlexically (as it can be the case of word stress assignment in different varieties of Italian, described in section 6.3 and referred to in note 57); or it may result from the role (not) played by certain constraints in a specific language (such as the Minimal Word requirement, mentioned in section 6.1). Our first task, carried out in chapter 5, is therefore to determine which phonological evidence may be relevant for the prosodic word in EP.

The application of the relevant tests for the assessment of the prosodic word domain will lead us to propose in chapters 5 and 6 an analysis of the way affixes, clitics and different kinds of words are prosodized. We adopt the view expressed in sections 6.3 to 6.5 above that a number of possible prosodizations are available – which may violate some aspects of the Strict Layer Hypothesis – and we assume the exclusion of the clitic group as a constituent of prosodic hierarchy.

Chapters 5 and 6 are divided according to the phonological properties of the constructions under observation. In the former, the constructions considered are those surfacing with a single word stress, a fact that will be shown to correlate with other phonological properties that are specific of these constructions. In the latter, the constructions analyzed are morpho-syntactic units that include more than one word stress. This property will also be shown to go together with various other phonological facts that distinguish these constructions both from the former type of constructions and from other structures formed by words grouped within a phonological phrase.

Chapter 7 is devoted to the study of reduction phenomena that affect clitic words, which are not regularly found within prosodic words.

Chapter 8 concludes the book with a review of our major findings, as well as a summary of the main contributions of this work for the theory of prosodic phonology and for the understanding of EP phonology.

Before we proceed, we survey in the following chapter the EP literature on the phonology of the word.

¹ Prosodic phonology is first developed as a reaction against Chomsky and Halle's (1968) conception that the syntactic surface structure and the phonological representation are related via a set of readjustment rules. These rules introduce boundary symbols into the surface syntactic structure. According to this approach, phonological rules refer to such boundary symbols in their structural descriptions, rather than directly to syntactic bracketing (see, for example, Selkirk 1980; Hayes 1989: 2.1.2).

² Although the syllable and the foot are not treated in earlier work (e.g. Selkirk 1980; Nespor and Vogel 1982), in Nespor and Vogel (1983, 1986) the prosodic hierarchy is already assumed to include these constituents. Itô (1986) and Zec (1988), cited in Inkelas (1990), propose, in addition to these constituents, the inclusion of the mora in the prosodic hierarchy. For Hayes (1989) the lowest constituent of prosodic hierarchy is the phonological word. Inkelas (1990), in turn, proposes the replacement of the mora, the syllable and the foot (considered to belong to the metrical structure) by other sublexical prosodic constituents, built with reference to the morphological structure. The main motivation for this view is to prevent lexical rules from referring directly to morphological structure. Selkirk (1986) also questions the status of the syllable and the foot as prosodic domains, although in later work (e.g. Selkirk 1996), the prosodic hierarchy is assumed to include these constituents.

³ In this book, we use the term prosodic word, instead of phonological word with which it is synonymous. We should point out that the latter was already at use before the development of prosodic phonology (see, among others, Câmara 1953, 1972; Dixon 1977).

⁴ Since the universal inventory and definition of prosodic domains is ultimately an empirical question, it is natural that some controversy in the literature is found. Thus, for example, on the basis of Greek data Condoravdi (1990) proposes a new prosodic category above the ϕ -level (see section 5 on an alternative proposal).

⁵ Languages like German and Dutch also show this type of behavior with respect to the phonological phrase: while there are no segmental rules that refer to this prosodic domain, there are other types of phonological phenomena that cue the ϕ -phrase (see Kleinhenz 1997 for relevant data and discussion).

⁶ The hierarchical nature of phonological representations is also a major departure from SPE linear conception, and it originates, in particular, with Liberman and Prince's (1977) proposal on English word and phrasal stress. Note that, under the latter approach, the hierarchical grouping of units is binary, and thus it creates a structure with unlimited depth. This property distinguishes metrical trees from prosodic trees, since the latter are flatter and their depth is limited (see, for example, Pierrehumbert and Beckman 1988: 6.1).

⁷ Early examples of prosodic structures that do not obey strict layering can be found in Nespor and Vogel (1983) and Nespor (1986). For example, in Nespor and

Vogel (1983) restructuring at the ϕ -level results in the creation of a ϕ' -node, defined in the following terms: a nonbranching ϕ which is the first complement of X on its recursive side loses its label and is joined to the ϕ containing X, under a new node labeled ϕ' (where X is a lexical head) (Nespor and Vogel 1983: 126).

⁸ “Lex” stands for a lexical head and “Lex^{max}” stands for the maximal projection of a lexical head. This formulation excludes maximal projections of functional categories from playing a role in the syntax-phonology mapping (see Truckenbrodt 1999 for a review and discussion).

⁹ Under Frota’s (2000) view, the weight conditions on I and ϕ -phrases play a role in the *basic* formation of these constituents, so that the mechanism of restructuring (Nespor and Vogel 1983, 1986) is no longer required. Notice, further, that the application of the weight conditions is constrained by the limits defined in (5) and (8).

¹⁰ The prosodization of clitic function words is not discussed in Frota (2000), who adopts basically the analysis proposed in Vigário (1999a), and developed in this book. Specifically, proclitic function words are assumed not to incorporate into the host prosodic words, and thus they count for the purposes of ϕ -branching.

¹¹ Although we do not have clear evidence for the foot domain in EP, the evidence for this domain in other languages argues for its inclusion in the prosodic tree. Since the prosodic tree is assumed to be universal, the foot should integrate EP prosodic trees as well. We should notice, additionally, that in example (11) we have avoided clitic elements, since one of the central questions addressed in this book concerns precisely the way these elements are prosodized. Prominence relations within prosodic constituents are not indicated (see §5).

¹² Frota’s contention is that this result is a derived one, since she claims that only in languages where focus is morphosyntactically marked is the prosodic structure directly affected by focus. Since EP has no such morphosyntactic properties, focus information does not affect prosodic phrasing in this language.

¹³ The following typology of prosodic processes may be drawn according to the way that phonological processes apply with respect to prosodic constituents (cf. Selkirk 1980; Nespor and Vogel 1986: 15):

Types of prosodic processes

- *Domain-span*: processes that apply *within* a given prosodic domain
- *Domain-limit*: processes that apply at the *edge* of a given prosodic domain
- *Domain-juncture*: processes that apply at the *edge* of a given prosodic domain, *within* a larger prosodic domain.

¹⁴ As Anderson (1992: 7.1) points out there are at least three notions of lexicon: (i) the locus of idiosyncrasy in language; (ii) the collection of all the items that belong to open or major word classes; and (iii) the scope of Lexical Phonology, as opposed to the domains of other processes characterized as postlexical. In this book we use the term *lexical entry* as the point where idiosyncratic information of

individual items is stored, and the terms *lexical component* or *lexical level* to refer to the point in the grammar where lexical processes apply.

¹⁵ Hannahs (1995a: 5.3), among others, suggests that the need for lexical levels can be correlated with the history of languages. Thus, languages where important parts of the lexicon are of different origins have been described as requiring the partitioning of the lexicon (as in the case of English, or Malayalam), whereas no such partition is necessary for languages with a lexicon of essentially a single origin (as in the case of French, or Italian – cf. Vogel 1991). As for the ordered conception of lexical levels, see Booij (1994: section 5), which also contains a review of important arguments against such a view.

¹⁶ The elimination of level ordering is not incompatible with the conception that lexical phonology precedes postlexical phonology, since the latter follows in a principled way from the organization of the grammar that is assumed (e.g. Booij 1996b).

¹⁷ As we will show in chapter 3, a number of lexical processes in EP are not structure preserving (e.g. vowel reduction – section 3; processes that create nasal vowels and glides – section 5; /e/-centralization – section 6; initial /r/-strengthening – section 9).

¹⁸ Precompiled phrasal rules have been proposed to account for similar effects in several other languages as well (e.g. Italian – cf. Nespor 1990, Peperkamp 1997a; Galician – cf. Fernández Rei 2002). Some facts of EP are also assumed to follow from precompilation in Vigário (1998b) and chapter 4, section 4.2.2 of this book.

¹⁹ For Kaisse (1990), there are two types of rules that may apply across words: P1 rules, which share a number of properties with lexical rules; and P2 rules, which have the properties considered below to characterize postlexical rules. According to Kaisse, both types of rules apply postlexically, P1 rules preceding P2 rules. Notice, nevertheless, that the existence of P1 rules does not preclude precompilation, as Kaisse remarks. In fact, if a rule that applies between words precedes (other) lexical processes, as is the case of vowel shortening in Kimatuumbi (Hayes 1990), then that rule cannot be postlexical – it must be a lexical one. In this book, we follow Hayes' proposal that if a rule applying between words has the properties of a lexical process, then it is a lexical process. In fact, as far as we can see, EP does not provide evidence for a split in the postlexical component.

²⁰ Hayes suggests, additionally, that precompiled rules are not sensitive to speech rate. However, Nespor (1990) holds that they may be sensitive to rate if the frame that defines the context for the application of precompiled rules includes a prosodic domain, as in the case of *Troncamento*.

²¹ Sensitivity to speech rate may be, at least partially, a consequence of the sensitivity of prosodic constituents to speech rate: if a process applies within a given prosodic domain, and if that domain may be smaller or larger depending on

the speech rate (see paragraph 7), then sensitivity of that process to speech rate may result from the variability of the prosodic domain limits.

²² A more or less widespread idea within lexical phonology is that function words, which do not belong to the major categories (N, V, and A), and do not enter in morphological operations, unlike affixes, are not present in the lexical component, and therefore they never undergo lexical rules (see the observations in Kaisse and Shaw 1985: 9, and, for an explicit claim along these lines, Inkelas 1990: 241). We show in chapter 5, section 4.2 that some EP facts contradict this view.

²³ See also sections 6.3 on the prosodization of clitics and 7 on resyllabification.

²⁴ Frota (2000) adopts Ladd's (1992, 1996) Compound Domain Hypothesis but she suggests that, at least for the intonational phrase level in EP, compound domains should be constrained in order to allow only for binary recursion.

²⁵ As shown in chapter 6, EP presents evidence for prosodic compounding at the prosodic word level as well. The analysis proposed there is extended to languages like English, Dutch, Italian, French and Baule (see also chapter 8).

²⁶ For Ladd (1992) only two conditions are needed to restrict the prosodic structure: the Ranking condition (which corresponds to Selkirk's Layeredness constraint), and the Uniformity condition, which is stated in (i).

(i) *Uniformity condition*

In a prosodic tree, all nodes immediately dominated by a given node N must be of uniform rank.

²⁷ Compound domains have been suggested to be the solution for some of the cases reported in the literature to weaken prosodic phonology theory and its underlying assumption that phonological processes do not refer directly to syntactic information. For example, under his Compound Domain Hypothesis, Ladd (1992) proposes to analyze Makonde's tone sandhi domain (presented in Odden 1990) as a compound domain, thus avoiding an approach with direct reference to syntactic information; and Frota (2000) suggests that compounding could also account for Greek data that led Condoravdi (1990) to postulate the existence of an additional prosodic domain ('Z') between the phonological phrase and the intonational phrase.

²⁸ French has been argued to have a stress domain larger than the word (cf. Pulgram 1970; Dell 1984, cited in Van der Hulst, Hendriks, and Van de Weijer 1999; Van der Leeuw 1997). On the contrary, Ladd (1996) claims that tonal distribution in French, as described by Dell (1984), can only be accounted for if it is assumed that the last full vowel of a prosodic word is metrically strong. Roca (1999) also analyses French, like other Romance languages, with a primary stress assigned at the word-level. This is, in fact, what is predicted by prosodic phonology theory – since it is assumed that each prosodic domain has a prominent element – independently of the possibility that phrasal stress might be more clearly perceived than word stress in some languages. Besides this, Hannahs (1995a,

1995b) presents segmental evidence for the prosodic word in French (e.g. glide formation).

²⁹ See Booij (1999), who shows that the relevant generalizations are constraints on prosodic rather than morphological constituents.

³⁰ In fact, we will show that in intonational languages, like EP, several types of tonal event also cue the prosodic word (see chapter 3, section 18).

³¹ Hall (1999a) notes, nonetheless, that there is one suffix in German (*-chen*) that can be deleted under identity, but does not show the behavior expected of a prosodic word, since it has a schwa as its only vowel.

³² Booij (1985) proposes in addition that reduction is restricted to occur only at the edges of phonological phrases, and that the two phonological phrases including the identical prosodic words must be adjacent.

³³ Notice that this type of operation is hypothesized in Zwicky (1977: 3) to be ruled out by a “Rule Immunity”, which states that “proper parts of words do not undergo rules of deletion under identity”.

³⁴ In section 5.1 we discuss in more detail the relevance of defining a minimal word size in European Portuguese, comparing it in particular with Italian. In fact, this language is similar to Portuguese in this respect, but has been claimed to have minimality requirements on the basis of the low frequency of monosyllabic words with open syllables (cf. Thornton 1996 and Peperkamp 1997a; see also Vogel 1993, and, for a different view, Bafille 1997).

³⁵ Wennerstrom (1993) considers that the presence of focus is also a test for the identification of prosodic words (a test also used in Hall 1999a for German). Raffelsiefen (1999a) argues against this position, since the possibility of focus may not correlate with other properties that can be established to identify English prosodic words. The notion of focus used by Wennerstrom seems parallel to the one we called initial emphatic stress in EP (see chapter 3, section 17). As we will see in chapter 5, in EP emphatic stress (but not focal stress) may be assigned both to *stressless* prefixes and to proclitics, which are independently shown not to form autonomous prosodic words.

³⁶ By contrast, Hayes (1989: 207) assumes that the prosodic word is always at least as large as the grammatical word. This view does not account for the existence of languages like Dutch where some suffixes behave like independent prosodic words.

³⁷ In Nespor and Vogel’s definition of the prosodic word domain of Type II languages, there is an additional clause that we have not included in (32), stating that unattached elements form independent prosodic words. This is meant to allow for elements that do not include stems, like function words, to form their own prosodic word. Under an analysis where these elements, being typically clitics, are not prosodic words, it is no longer necessary to include this information in the definition of the PW domain (see next section, on the prosodization of clitics).

³⁸ See Peperkamp (1997a: 5.4.1.2) for the discussion of an alternative analysis of Neapolitan facts whereby the verb and the clitic each form an independent prosodic word.

³⁹ The prosodization of Standard Italian enclitics is discussed in chapter 8 in the light of our proposal on the prosodization of EP enclitics.

⁴⁰ In yet other proposals, such as Berendsen (1986), clitics are conceived as floating elements whose prosodization may freely vary. For example, in Dutch they would optionally incorporate into the prosodic word, the phonological phrase, or both.

⁴¹ The arguments for the proposal that in the most frequent case enclitics in German are incorporated into the preceding prosodic word are weaker in the sense that the data is consistent with, but does not force, such an analysis. This proposal contrasts with Kleinhenz's (1996) analysis of German enclitics, according to which these elements are adjoined to the preceding prosodic word.

⁴² In a theory such as Inkelas' (1990), the direction of cliticization follows from the prosodic subcategorization frame of clitics. In this case, therefore, the direction of cliticization is always a property of individual clitics. Such a theory does not account for the optionality in the direction of cliticization found, for example, in Dutch (see also Van der Leeuw 1997).

⁴³ This is proposed to be the case of all European languages surveyed in Nespor (1999b), including, besides Greek and Italian, Dutch, Icelandic, Spanish, French, Romanian, and Portuguese.

⁴⁴ Nespor (1999b) claims that word+word compounds are always mapped onto two prosodic words, a hypothesis supported by many Romance, Germanic, and Slavic languages.

⁴⁵ Other languages where stem+word compounds are argued to be mapped onto two prosodic words are Catalan, Romanian, and Serbo-Croatian (Nespor 1999b).

⁴⁶ In order to derive the latter possibility, Peperkamp (1997a: 127) proposes a constraint that allows for familiar compounds to be mapped onto a single prosodic word.

⁴⁷ See also Di Sciullo and Williams (1987), who propose that noun+noun compounds in French are in fact listed phrases, that is, lexicalized syntactic phrases.

⁴⁸ There are, nevertheless, two such suffixes that do not form independent prosodic words in the language.

⁴⁹ As shown in chapter 5 and section 6.4 of this chapter, EP, like other Romance languages, patterns differently from Dutch and (most varieties of) German in that the segmental make-up of suffixes does not correlate with the way these units are prosodized. For instance, the great majority of consonant-initial suffixes are incorporated into the prosodic word that dominates their morphological base.

⁵⁰ Notice, nevertheless, that Raffelsiefen (1999a: 143) shows that unproductive prefixes in English and in German can form independent prosodic words.

⁵¹ Peperkamp (1997a: 30) points out that there seems to be a correlation between the (im)possibility of phrasal resyllabification, and the type of syllables found in a given language. Romance languages, Turkish and Korean show simple syllable structure and allow for phrasal resyllabification, whereas Germanic languages show complex syllable structure and do not allow for phrasal resyllabification.

⁵² The term resyllabification may also be used to refer to the cyclic operation of syllabification that applies word internally (e.g. Nespor and Vogel 1986: 106, Rice 1989). Here, we are only concerned with resyllabification across word-boundaries. The need for the distinction between (lexical) syllabification and (postlexical) resyllabification is demonstrated, for example, in Nespor and Vogel (1986: 68–72), on the basis of Spanish and French data.

⁵³ In the examples, we omit the details concerning the syllable internal structure for ease of exposition.

⁵⁴ The data provided by Dutch (or German, see for example, Wiese 1996; Kleinhenz 1996; Hall 1999a) involving vowel-initial clitics further show that, like in Romance languages, in Germanic languages (postlexical) resyllabification occurs, although it has a lower domain – the prosodic word.

⁵⁵ This proposal is meant to account for resyllabification both between words and between prefixes (whether stressed or unstressed) and their morphological bases. In fact, the prosodic word boundary readjustment allows Peperkamp (1997a) to analyze Italian prefixes ending in a consonant and followed by a base starting with a vowel not necessarily as elements incorporated into the base prosodic word. As we have seen in section 6.2, the latter analysis is proposed in Nespor and Vogel (1986) for Italian prefixes ending in consonant, regardless of the presence of a context for resyllabification.

⁵⁶ Researchers are not always aware of the importance of isolating this variable. In EP, the observation of more than one dialect in the description of a given process is possibly the most common practice. In other languages, such as German, we have seen that the prosodization of affixes may vary from dialect to dialect (cf. section 6.5). The same may happen with clitics (as is shown by Serbo-Croatian and Italian dialects – cf. section 6.3). Notice, nevertheless, that it can also happen that data from different dialects converge towards the same analysis, as is suggested to be the case of German clitics in Kleinhenz (1996). In any event, this source of variation should not be discarded.

⁵⁷ An illustrative example of how this point may be crucial for a given analysis is provided by the account of the prosodization of pronominal clitics in Standard Italian, developed in Peperkamp (1997a) (see section 6.3). In this variety of Italian, the presence of enclitics does not affect the assignment of word stress: word stress falls only on the host, in the same place as when it occurs in isolation. This is taken to indicate that the enclitic occupies a position in the prosodic tree external to the prosodic word. Nevertheless, if word stress assignment is not operative postlexically, this behavior is in fact explained independently of the way postlexically combined elements are prosodized – the process is simply not active

at the moment the relevant elements occur together. EP behaves similarly to Standard Italian in that pronominal clitics do not interact with word stress. Interestingly, this language provides evidence to choose between an analysis of this fact as following from the prosodization of enclitics, or as following from the non-application of word stress assignment within the postlexical component (see sections 3.1, 5.3.3.1, and chapter 8 for the discussion of this issue).

Chapter 2

Previous studies on European Portuguese word phonology

1. Introduction

In this chapter we survey some of the most salient aspects found in the phonetic/phonological literature on EP that may be related either to the phonological word or to clitics. To our knowledge, a comprehensive study on these topics has not been previously conducted on this language. Consequently, most of the facts presented here appear to be dispersed in the work of different researchers.

2. Pregenerative studies

Among the most prominent pregenerative researchers that have described EP are Gonçalves Viana, Sá Nogueira, Câmara, and Morais Barbosa. In the following paragraphs we present aspects of the work of each of these linguists that may bear a connection with the issues treated in this book.

One of the earliest phonetic descriptions of what may be called Modern European Portuguese was conducted by Gonçalves Viana at the end of the 19th century. Gonçalves Viana (1883) provides an extremely accurate description of the realization of European Portuguese as spoken in the Lisbon region. Although the language described is not precisely the same as present-day EP spoken in the same region, the similarities are remarkable. In fact, much of the phonological phenomena found in more recent work, such as Morais Barbosa (1965), Mateus (1975), d'Andrade (1977), and Mateus and d'Andrade (2000), are already reported in Gonçalves Viana's work.

Despite the absence of prosodic structure, the relevance of some sort of *domain* for the observation of phonetic facts is implicit in Gonçalves Viana (1883). This may be deduced from the structure of his presentation. First, the facts are described considering segments within the syllable, then within the word, and finally across words.

Gonçalves Viana uses the term *word* in its morphological sense. This unit is seen to be formed by syllables, whose make-up may be dependent on

their position. For example, (i) a syllable may contain a nasal diphthong, but not if it is word initial or word internal, (ii) a syllable may start with a “simple” *r*, but not in word initial position, and (iii) vowels in syllables closed by *r* may reduce, but only in word initial or word internal position. Word initial position is also reported to induce the realization of palatal vowels as high, and to favor the realization of *a* and *o* in closed syllables as low. As we will see in chapters 3 and 5, most of these facts may be seen to result from the application of phonological processes that refer to a *prosodic domain*, specifically, the prosodic word.

When considering the facts across words, Gonçalves Viana observes in a systematic way sequences of items consisting of a stressed word and an unstressed function word, and only rarely sequences of stressed words. After comparing sequences of vowels across words with parallel sequences within the word, Gonçalves Viana (1883: 39) concludes that, in general, the same rules apply in both conditions. Nevertheless, since in his description of across word phenomena one of the (morphological) words involved is usually a clitic, what seems to be captured by Gonçalves Viana is the phonological coherence of clitics with respect to their hosts, rather than the behavior of sequences of (phonological) words.

Despite the similarity between the combination host+clitic/clitic+host and single words, a few differences can, nevertheless, be extracted from Gonçalves Viana’s data: a stressless [i] in word initial position does not surface as a glide if preceded by a vowel, while it does so within a single word (see 47a); and the alternation [j]/0 when a stressless *e* is followed by a vowel is only possible across words, but not within words, where the deletion of *e* (0) is impossible (see 47b).

- (47) a. *a igreja* [i] versus *eira* [j]
 ‘the church’ ‘threshing floor’
 b. *de ouro* [j]/0 *recear* [j]
 ‘of gold’ ‘(to) fear’

From various remarks, we gather that Gonçalves Viana considers the conjunction *e* ‘and’, the definite articles *o* ‘the-MASC’ and *a* ‘the-FEM’, pronominal clitics, such as *me* ‘to-me’ and *lhe* ‘to-him/her’, all the prepositions, including *de* ‘of’ and *para* ‘to’, the quantifier *cada* ‘each’, and the complementizer/pronoun *que* ‘that/which’ unstressed. These remarks include (i) a comparison between the behavior of vowels in monosyllabic function words with the behavior of stressless vowels within the word in

An exhaustive list of stressless elements is, nonetheless, not provided, and in some cases, there is some overgeneralization. For example, it is not the case that in EP all prepositions are unstressed, since forms such as *apÓs* ‘after’ or *sObre* ‘on’ bear word stress. This is shown by the fact that the stressed vowels of these words do not undergo vowel reduction, unlike, for instance *de* and *para* (see chapter 5 for the identification of clitic words in EP, and chapter 6, section 5.1 for a discussion on the sources of the asymmetry between stressed and unstressed units).

(48) a. *pOrta-machAdo* (but *t[u]rcicOlo*)
 ‘ax bearer’ ‘torticollis’
 b. *sEcamEnte*
 ‘crudely’
 c. *prEgozInho*
 ‘little nail’
 d. *recomendÁ-lo-Ia*
 ‘(I) would recommend it’

The classes identified by Gonçalves Viana to exhibit two primary stresses (or, in a different terminology, a strong secondary stress and a primary stress⁵⁹) are the same as those usually considered in more recent discussions on word stress. We should notice, nevertheless, that there are other constructions that show more than one word stress, namely non-lexicalized morphological compounds and derived words with stressed prefixes (e.g. Pereira 1999; Vigário 2000a, and chapter 6 of this book).

Sá Nogueira (1938) presents another detailed phonetic description of EP, where a comparison with Gonçalves Viana's observations is usually drawn. Sá Nogueira also investigates the realization of adjacent vowels both within and across words. However, unlike Gonçalves Viana, he looks

at stressed words instead of unstressed function words in his observation of across word phenomena. Additionally, he is more exhaustive in his study, since he considers all possible combinations of stressed and unstressed vowels. Despite this, the absence of an explicit comparison between the phonological behavior of adjacent vowels within and across words renders Sá Nogueira's work less interesting from our point of view. Moreover, no additional remarks are made that could bear on the relevance of the (phonological) word for the phonetic description of EP.

In the structuralist work of Câmara (1953, 1972), the variety of Portuguese systematically described is the one spoken in Brazil.⁶⁰ However, we can find several references to the variety spoken in Portugal, together with an explicit distinction between phonological objects and morphological units. This justifies the inclusion in this brief review of a linguist who does not specifically address the European variety of Portuguese.

Câmara (1972) is one of the few works on Portuguese where major role is assigned to the phonological word. Explicitly drawing the attention to the distinction between the phonological word and the morphological word ("formal unit"), he presents two paradigmatic examples of this lack of correspondence: the case of clitics that attach phonologically to a free form, as exemplified in (49), where two formal units correspond to a single phonological word; and the case of juxtaposed compounds, exemplified in (50), where two phonological words correspond to a single formal unit.

- (49) *feriu-se*
 '(he) has hurt himself'
 o belo
 'the beauty'

- (50) *guarda-chuva*
 keeps rain
 'umbrella'

Câmara shows a remarkable precision when dealing with the delimitative function of word stress. Since word stress location in Portuguese is to some extent variable, as it may occur in any of the three final syllables of a word, word stress location *per se* does not signal word boundaries. By contrast, in Brazilian Portuguese (BP) the unstressed vowel system may cue the word boundary, for in this variety pretonic vowels are less reduced than posttonic ones. Thus, BP pretonic and posttonic vowel system may signal phonological word boundaries. We should notice,

nonetheless, that this is not the case of European Portuguese, since in this variety both pretonic and posttonic vowels are, in general, equally unstressed/reduced.

Considering the specificity of clitics, and observing in general the Brazilian variety of Portuguese, Câmara claims that clitics become part of the phonological word to which they attach. When they are proclitic, their vowels are (in general) not distinct from pretonic ones, whereas when they are enclitic their vowels behave like posttonic vowels. However, proclitic vowels may also behave like the unstressed final vowels of the clitic itself (Câmara 1972: 35).⁶¹

Without explicitly precluding enclitics from receiving stress, he assumes that proclitics may form an independent prosodic word under “emphasis”, both in Brazilian Portuguese and in European Portuguese. Emphasis on clitics is taken to create “certain anomalies” in the stressed vowel system, since vowels that otherwise do not appear in stressed position in EP, namely [i] and [ɐ] (in non-nasal environments), do get stressed (Câmara 1972: 72). Also within the spirit of the “anomalies” produced within the clitic system, and considering in this case the Brazilian variety, this linguist notes that the clitic word *mas* ‘but’ may be realized with a mid vowel – instead of a low [a] – which is in general not allowed in stressed position.

As we will see in chapters 3 and 5, both types of anomalies can be accounted for if we assume that some of these phenomena are lexical (e.g. vowel reduction and word stress assignment), while others are postlexical (e.g. the reduction of clitic words and emphatic stress).

Morais Barbosa (1965) offers a structuralist analysis of EP phonology, where phonological segments are systematically described according to their position within the word. The notion of word used throughout this work is phonological rather than morphological. Nevertheless, apart from some facts related to word stress, no phonological distinction is established between phonological and morphological words.

Observing pronominal clitics and other function words separately from words that bear their own stress, Moraes Barbosa concludes that the vowel system of enclitics (and also of proclitics) is the same as the word final unstressed vowel system. Moraes Barbosa assumes that postverbal pronominal clitics are enclitic and preverbal pronominal clitics, as well as other function words, are proclitic.⁶² However, the behavior of enclitics and proclitics with respect to their hosts is not systematically described.

Apart from pronominal clitics, other function words that are considered to be clitic are presented in a non-exhaustive list, which includes articles,

prepositions, complementizers, and conjunctions. Like in Câmara (1953, 1972), some of these words are reported to become non-clitic under stress.

Morais Barbosa (1965) also points out that in some forms there is no correspondence between the phonological and the morphological notions of word. These cases are those already identified by Gonçalves Viana to bear two word stresses – the derived words with suffixes such as *-zinho*, *-zinha*, and adverbs formed with *-mente*, which are paralleled to word compounds, where two word stresses are also found.

3. SPE studies

Mateus (1975) and d'Andrade (1977) are two landmarks in standard generative analyses of EP phonology. In both of these studies, as well as in Andrade (1980) and d'Andrade (1994a, 1994b), also developed within Chomsky and Halle' (1968) SPE model, the term *word* is always used in a morphological sense.

In her analysis of word stress, Mateus (1975) observes nouns, verbs and adjectives, and proposes three general rules, plus one rule specific of verbs. With respect to words that are usually considered to bear more than one *primary* stress, such as those with the evaluative suffixes starting with a [z], and adverbs with *-mente*, Mateus proposes that they undergo a special derivation, which is signaled by a special “derivative boundary” (=).⁶³ Stress assignment is cyclic and thus the presence of stress in the rightmost element of these words is obtained. The occurrence of a non-reduced vowel in the primitive word is seen to follow from the presence of the derivative boundary, which triggers the blocking of a rule that weakens stressless vowels (cf. Mateus 1975: 4.2). The internal structure assigned to these forms is illustrated in (51).⁶⁴

- (51) *bolazinha* b[ɔ]la=zInha
 ‘little ball’
 belamente b[ɛ]la=mENte
 ‘beautifully’

Under this view, the unreduced vowel of the base is not considered to bear a primary stress, unlike is usually assumed in EP literature, at least at some point of the phonological derivation.

D'Andrade (1994b) is also concerned with the words formed with the suffixes starting with a [z]. In contrast with Mateus (1975), d'Andrade argues against the presence of a = boundary between the base and this suffix. Specifically, d'Andrade (1994b) follows a previous analysis (d'Andrade 1977: 6.16), where this boundary is proposed to be crucially involved in the non-application of a rule of stop spirantization. In a form such as *opaquito* (opa[k]=ito) 'opaque-DIM' the rule does not apply because there is a = boundary, whereas it applies in similar phonological environments if a different boundary follows the relevant consonant, as in *opacidade* (opa[s]+idade) 'opacity'. Since the vowel that precedes = is usually reduced (cf. op[ɐ]quito), d'Andrade concludes that the same boundary may not be involved in the forms with a [z]-suffix, where the vowel that bears stress in the primitive word always surfaces unreduced. This author therefore assumes that the boundary that separates the primitive word and the [z]-suffix must be a word boundary (#). The presence of inflectional affixes before this suffix, as documented in (52), supports the word-boundary analysis.

- | | | |
|------|--|--|
| (52) | <i>cão / cães</i>
'dog / dogs' | <i>cãozito / cãesitos</i>
'little dog / little dogs' |
| | <i>vermelho / vermelha</i>
'red-MASC / red-FEM' | <i>vermelhozinho / vermelhazinha</i>
'little red-MASC / little red-FEM' |

For d'Andrade, these word-forms are thus parallel to true compounds, and the absence of vowel reduction follows from the assignment of stress to each word (d'Andrade 1994b: 46). That the base this suffix attaches to corresponds to a morphological word will remain a common assumption in subsequent work (e.g. Mateus 1983; Villalva 1994; Pereira 1999; Mateus and d'Andrade 2000).

Andrade (1980) is a preliminary study on issues related to vowel reduction, which is developed to some extent along the lines of proposals in Mateus (1975) and d'Andrade (1977). Of particular interest to us is a short remark on the asymmetry between word initial and word final positions. The observation that vowel neutralization is less extreme in initial position leads Andrade to put forth the hypothesis that the beginning of the word is perceptually more important than the end in Portuguese (Andrade 1980: 65).

Although syllable structure has been claimed to play a major role in the reduction of initial vowels (cf. Mateus 1995; Mateus and d'Andrade 2000), similarly to Andrade, we also argue that word initial position is crucial to

account for the less reduced properties of initial vowels, and we add empirical support for the importance of this position for a number of other phenomena in EP (see chapter 3, section 10). We further offer arguments in favor of considering that the relevant domain is phonological rather than morphological.

In the work referred to above, we can find many rules that include a word boundary in their structural description. Some of these rules are reviewed in chapter 3, where we discuss the morphological/phonological nature of their domain of application.

4. Post-SPE studies

4.1. Words with two primary stresses

In Mateus (1983), a new analysis of stress is proposed where morphological information plays a major role. The relevance of morphological information for stress assignment will be adopted in most subsequent work on word stress in EP (e.g. d'Andrade 1988; d'Andrade and Laks 1992; Van der Leeuw 1997; Pereira 1999; Mateus and d'Andrade 2000).

Essentially, Mateus (1983) proposes that stress is assigned to the last vowel of the stem in the case of nouns and adjectives, and to the last vowel of the theme in the case of verbs. As for the presence of two word stresses in future and conditional verb forms with mesoclis (i.e. with internal clitics), and in derived words with evaluative suffixes starting with *-z-*, and with *-mente*, it is claimed to result from the presence of a (morphological) word boundary intervening between the base ("primitive word") and the suffix. Such a boundary allows for the preservation of the primitive stress. The examples in (53) show the structure assigned in Mateus (1983) to the constructions with internal clitics and with suffixes starting with *-z-*.

- (53) *bater-nos-íamos* ## bat e + r # nos # + ia + mos ##
 '(we) would hit each other'
 bolazinha ## bəl + a # ziŋ + a ##
 'little ball'

The classes of units with a less clear morphological structure, such as conjunctions, prepositions, pronouns, and prefixes, are, nevertheless, not discussed in Mateus (1983), as well as in most recent work. These units are problematic because some elements that belong to a given class bear word

stress, while others do not, as shown in (54) (see also chapter 6). In addition to this, given the relevance of morphological notions such as *stem* and *theme* for word stress assignment, it remains unclear how the elements belonging to these closed classes may be assigned word stress.

(54)

	STRESSLESS		STRESSED	
CONJUNCTIONS	mas	‘but’	porém	‘but’
PREPOSITIONS	para	‘to’	após	‘after’
PRONOUNS	te	‘you’	tu	‘you’
PREFIXES	des–	‘un–’	pós–	‘post–’

Still in the realm of word stress assignment, d’Andrade (1988) analyzes the same structures within metrical theory. The difference between regular suffixes and *–mente* or *z*-suffixes is seen to result from the fact that the former attach to a derivational stem (“radical derivacional”), whereas the latter attach to the word. It is further argued that it is because there is a word inside another word that the stress in the primitive word is preserved (d’Andrade 1988: 22).

The procedure for metrical tree construction indicates that these suffixes are treated as words themselves, and are autonomous for the purpose of tree formation until the “last tree” is built (d’Andrade 1988: 22). In the last tree, which joins the two non-terminal trees, the last level of stress is assigned to the rightmost strong element. This accounts for the fact that the stronger stress of these constructions corresponds to the second stressed unit.

As in Mateus (1983), the presence of stress in these suffixes remains somewhat unclear. The domain of stress is assumed to be the *word*, defined by the frame $[...]_{\alpha}$ (where α stands for verb, noun, adjective or adverb). Nevertheless, these suffixes *per se* do not belong to any of these categories.⁶⁵ Moreover, the categories of words defined as domains for stress assignment exclude prepositions, conjunctions or prefixes, and thus the presence of stress associated to such units is not accounted for.

With respect to mesoclititic structures, d’Andrade (1988) departs from Mateus (1983) by assuming that clitic pronouns are infixed to the inflected verbal forms in future and conditional tenses. Future and conditional verbs are assumed to have a derivational stem distinct from other verb forms, which accounts for their specificity in terms of stress location. As for the stress pattern found in the verbal base of mesoclititic structures, it is argued to follow from the presence of the infixed pronoun: because the pronoun is taken to be extrametrical, the prominent element of the preceding foot

corresponds, by convention, to the vowel that immediately precedes the pronoun, which is the theme vowel when a single pronoun is infixed.⁶⁶

We will disregard the details of d'Andrade's analysis of stress assignment in mesoclitic constructions, for we would like to focus our attention on two aspects of his underlying conception of these constructions. First, the pronoun is considered to be infixed to an inflected verbal form. Nevertheless, from the proposed analysis no motivation emerges either for this possibility, or for the impossibility of finding pronouns inside other verbal forms. Second, under this approach the verb forms inflected in the future and conditional are predicted to be underlyingly identical whether the pronoun is infixed or not. However, as we will see in chapter 4, there are some cases where the two forms do not coincide.

Van der Leeuw (1997) proposes an analysis of EP mesocclisis in terms of Optimality Theory. This proposal is a step forward in the analysis of mesocclisis since it is intended to capture some basic facts related to this construction: the presence of internal clitics only with these verbal forms; the specific position of clitics in these constructions; and the stress pattern of mesocclisis.

Assuming the bipartition lexical/postlexical component, Van der Leeuw claims that postlexical candidates for evaluation may be generated with two primary stresses. The presence of two primary stresses implies the existence of two prosodic words, since a prosodic word is required to contain one and only one primary stress. In structures with clitics, the candidates with two prosodic words are then showed to give better results in terms of the proposed constraints' evaluation, than those with a single prosodic word, and thus the winning candidate presents two prosodic words.⁶⁷

The presence of internal clitics, on the other hand, is assumed to occur only with the future and the conditional verb forms because only these verb forms obey the restriction of *Morpheme Ordering Constraint*, which states that a stem must be followed by a tense morpheme, and a person morpheme must be preceded by a tense morpheme. Therefore, in other verb forms, which lack two adjacent tense morphemes, no internal clitic is allowed.

Finally, the possibility of internal clitics is considered to be a case of encliticization to the first prosodic word of the verb form, and the preference for the internal cliticization over encliticization to the whole verb form is seen to be a consequence of a constraint on stress location, the *Stem Dominance Constraint* (SDC). This constraint states that stress falls on the verb stem. As said above, the presence of internal clitics favors the selection of a candidate with two prosodic words, and the existence of two prosodic words allows stress to fall on the first prosodic word and, thus, to

obey SDC. Van der Leeuw then assigns a low ranked status to the constraint *STRUC (“avoid structure”), argued to play a role in Portuguese because, in constructions with no clitic, it is responsible for the preference for structures with a single prosodic word instead of two (e.g. [mataríamos] instead of *[matár][íamos] ‘(we) would kill’).

Putting aside the actual explanatory power of this analysis, which we will not discuss here, we believe that the same general comment made on d’Andrade’s (1988) conception of mesoclisism applies here. As it stands, this analysis does not account for possible cases of mismatch between lexical inflected forms and mesoclitic constructions, since, apart from the presence of clitics, the prosodic structure and stress, the candidates that are selected as optimal outputs are similar to the lexical output forms. As said above, there are a few forms where the verb stem in mesoclisism does not coincide with the verb stem selected when no internal clitic occurs, thus suggesting that the two formations are obtained independently of each other (cf. chapter 4, section 4.3).

Besides this, it remains unclear why a clitic should be allowed to appear between two tense morphemes, but not between a tense and a person morpheme, or between a verb stem and a tense morpheme. Since clitics usually appear in the periphery of inflected words, clitics should in fact be prevented from appearing inside proper inflectional affixes. An alternative view is to consider pronominal clitics to be inflectional affixes themselves. In this case, however, the bulk of phenomena that distinguish inflection from cliticization in EP, which we present in chapter 4, remains unexplained. Notice finally that, even if clitics are considered to be inflectional affixes, the relative ordering of clitics and tense morphemes remains a puzzle. In fact, on the basis of the observation of more than 60 languages, Bybee (1985: 33–35) concludes that inflectional morphemes are usually ordered according to their meaning, following the hierarchy of *aspect-tense-mood-person*. Thus, given that pronominal clitics cannot be analyzed as tense morphemes, they should not be allowed to intervene between inflectional affixes of this type (see chapter 4 for a more detailed discussion on the affix/clitic status of stressless pronouns in EP).

A somewhat different analysis of future and conditional forms is found in Mateus and d’Andrade (2000), where these verb forms are related to the future periphrastic construction *haver de* + infinitive verb that originated them (e.g. *hei-de falar* ‘(I) will speak’). The future is considered to be formed by an infinitive verb plus tense and person markers which are assumed to be the present indicative form of the auxiliary verb *haver*, and the conditional is taken to be composed by the infinitive verb followed by

52 *Previous Studies*

haver inflected in the imperfect tense.⁶⁸ Because of the position of pronominal clitics, these verb forms are seen to be exceptional, since clitics may appear internally, between the verb formative identical to the infinitive and the specific endings of the future and the conditional forms.

Although we also believe that there are good reasons to analyze the units with internal clitics as periphrastic constructions, our analysis differs in some crucial aspects from Mateus and d'Andrade's approach. The most important one is that we defend the idea that only the constructions with internal clitics are obtained via the insertion of an infinitive verb form and a(n auxiliary-like) affix, while the future and conditional verb forms without internal clitics are inserted lexically inflected at the moment of phonological instantiation (see chapter 4, section 4.3).

Pereira (1999) is the most comprehensive study on word stress in EP we could find. This analysis is developed within the metrical theory framework, in the version first presented in Idsardi (1992). Pereira adopts the lexical phonology view that the grammar has two distinct components: lexical and postlexical. Since several aspects argued to be relevant for word stress assignment are proper of the lexical component, this is explicitly seen to be a lexical operation.⁶⁹ These aspects include (i) the relevance of morphological domains for word stress assignment, along the lines of Mateus (1983)⁷⁰, (ii) the lexical marking of certain simple roots and affixes to bear stress; (iii) the lexical specification of certain roots and suffixes not to be subject to some specific parameters proposed; (iv) the lexical specification that certain roots are subject to special parameter settings.

Following Vigário (1999b, 1999c) (developed in chapter 4 of this book), Pereira considers cliticization to be a postlexical operation, and therefore clitics are not taken into consideration in the proposed parameter settings for verbal stress assignment, which is accomplished within the lexical component. As for the stressless nature of clitics, it is assumed to be part of their lexical specification.

Also adopting Vigário's (1999b, 1999c) proposal on mesoclisism (see also chapter 4 of this book), Pereira assumes this construction to be obtained postlexically, through the concatenation of units that are assigned stress independently within the lexical component.

The other constructions exhibiting two primary stresses are also considered to be the result of postlexical operations. These constructions include (i) syntactic compounds (formed by the concatenation of words), (ii) certain morphological compounds (those argued to be formed by the concatenation of a stem with a word, along the lines of Nespor 1999b), (iii) prefixed words with stressed prefixes, and (iv) derived words with *-mente*

and *z*-evaluative suffixes, which attach to (inflected) words. In all these constructions, stress is assigned to each constituent in the lexicon, and follows either from the general parameter settings that derive word stress, or from information stored in the lexical entry of the relevant constituent.

Conversely, unstressed prefixation and suffixation, as well as root compounding are assumed to be accomplished in the lexicon.

For Pereira (1999), the assignment of morphological processes to the postlexical level is a consequence of the locus of the construction of prosodic domains. Considering the cases where more than one prosodic word corresponds to a single derived or compound word, Pereira assumes that the relevant prosodic words are joined together in a constituent higher than the prosodic word – the phonological phrase. It is because this domain is built postlexically that the concatenation of the relevant units that form independent prosodic words is argued to occur postlexically.⁷¹ As we will see in chapter 6, there is good evidence that most of the units assumed by Pereira to form phonological phrases, constitute a compound prosodic word, instead. Nevertheless, since compound prosodic words are also postlexically obtained (see chapter 6, section 5.3), our findings do not undermine her general idea.

4.2. Cliticization

Compared to word stress, the issues related to the cliticization of stressless words in EP have received much less attention, and are often reduced to short notes and are referred to as side issues.

In her proposal about the construction of phonological domains, Viana (1987) assumes pronominal clitics to be already part of the prosodic word when intonational units are built (Viana 1987: 112). However, Viana does not elaborate on the notions of part of and prosodic word, since her goal is to account for the prosodic domains relevant for tonal association, the lower of which is argued to be the *intonational unit* (a unit that is assumed to be parallel to the phonological phrase of Selkirk 1984 and Nespor and Vogel 1983 or the “*groupe accentuel*” of Verluyten 1982 – cf. Viana 1987: note 3.3).

The units belonging to the categories Noun, Verb, Adjective, and Adverb constitute the heads of intonational units. As for the grammatical words that do not belong to any of these categories, they are included in the same intonational unit as the following lexical head. A distinction is, therefore, implicitly drawn between clitic pronouns and, for example, definite

articles or prepositions – the former attach at the level of the prosodic word and the latter attach at the level of (a constituent similar to) the phonological phrase. Again, this is not the topic under research in Viana (1987) and thus the subject is not further developed. We should notice, nonetheless, that according to the facts discussed in chapter 5, there seems to be no phonological reason for distinguishing EP clitic pronouns from other clitic function words, such as articles or prepositions.

Some of the proposals related to pronominal clitics in EP bear more specifically on aspects related to the direction of cliticization. This is the case of Brandão de Carvalho (1989), Frota (1994), Galves and Galves (1995), Cameira (1994), Barbosa (1996), Van der Leeuw (1997) and Kaiser (1999), among others.

In contrast to what is assumed by other EP linguists (e.g. Gonçalves Viana 1883; Morais Barbosa 1965; Barbosa 1996), Brandão de Carvalho (1989) defends the idea that pronominal clitics, either in preverbal or in postverbal position are always phonological enclitics, and thus differ from other “particles”, such as *que* ‘that’, *se* ‘if’, or *de* ‘of’, which are always proclitic. A basic argument for this proposal is pronominal clitic distribution. Specifically, the impossibility of preverbal clitics to occur in sentence initial position is claimed to result from the obligatory enclitic nature of stressless pronouns. Since these clitics are assumed to lean on a preceding host they cannot appear in a configuration where no such host exists.

Behind this idea is the belief that syntactic proclisis can occur not only with *proclisis triggers* (certain adverbs and quantifiers, interrogative and negative words, and complementizers), but also with any word preceding the clitic. This is illustrated in (55), taken from Brandão de Carvalho (1989: 409).

- (55) *Eu-te vi* ‘I saw you’
 I-you saw
 Ele-te disse ‘He said (it) to you’
 He-you told
 Eu-te vou dizer ‘I will tell you’
 I-you go tell

The author finds syntactic proclisis in these cases marked, although possible. However, proclisis in these contexts never occurs in EP. Indeed, it is generally acknowledged in the syntactic literature that only syntactic enclisis, but not proclisis, is allowed in contexts where the clitic is not preceded by a proclisis trigger (see, for example, Duarte and Matos 2000

for a review). This means that the same facts that ban syntactic proclisis in forms like those in (55) may be considered to be responsible for the ban on preverbal clitic positioning in sentence initial position, that is, the lack of a proclisis trigger.

There are a number of facts that seem to challenge the enclitic nature of preverbal pronominal clitics in EP. First, the possibility of preverbal pronominal clitics to be preceded by another clitic in sentence initial position (e.g. a stressless complementizer) suggests that no prosodic word is actually required to the left of the clitic pronoun (but see Brandão de Carvalho 1989: note 4). In addition, there are contexts where initial clitics may be found in EP, as when a parenthetical expression is inserted between the trigger and the clitic, as in (56).

- (56) *O João disse que, se tudo corresse bem, o viriam buscar cedo.*
 the John said that if everything went well him (they) would come
 pick early
 'John said that, if everything went well, they would come pick him
 up early.'

Finally, and most importantly, many phonological facts, which are presented in chapter 5, contradict Brandão de Carvalho's hypothesis, since they show that pronominal clitics behave phonologically like proclitic elements when preverbal and as enclitic units when postverbal. Furthermore, preverbal pronominal clitics pattern with other proclitic words, as far as those facts are concerned. (See also chapter 7, section 5.4.1 for our account of why pronominal clitics are excluded from sentence initial position in EP).

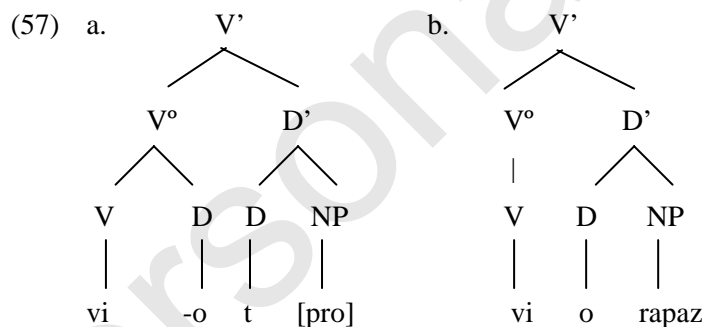
Besides Brandão de Carvalho (1989), Galves and Galves (1995) and Kaiser (1999) also assume the enclitic nature of stressless pronouns in EP, whether preverbal or postverbal.⁷² In fact, the idea that preverbal pronominal clitics are enclitic to the word that precedes the verb is already found in Said Ali (1964). However, for Said Ali enclisis is restricted to occur only if the preceding word is unstressed, as is the case of complementizers, or if it is strongly stressed, as it is considered to be the case of Wh-words and emphatic words.

The hypothesis that pronominal clitics are specified to phonologically attach to the proclisis trigger is also explored in Frota (1994), Frota and Vigário (1996) and Vigário and Frota (1998). In these works, proclisis triggers are phonologically characterized as *strong* function words. Here, however, the phonological properties of enclisis and proclisis in EP are seen to indicate that preverbal pronominal clitics are not enclitic. This fact,

56 Previous Studies

as well as the possibility of preverbal clitics to appear non-adjacent to the proclisis trigger, has led Vigário and Frota (1998) to propose that the syntactic requirement that pronominal clitics attach to the verb creates a conflict between syntactic and phonological information. This conflict is always solved on the behalf of the syntax. Thus, in the mapping between syntax and phonology, the syntactic configuration of proclisis imposes a phonological prosodization of pronouns that results in phonological proclisis instead of enclisis. A mismatch between the phonological requirements of pronominal cliticization and the actual prosodization of clitics therefore emerges.

Another attempt to characterize the special behavior of clitics is Cameira (1994). Starting by pointing out the similarities between postverbal accusative pronominal clitics and definite articles, and assuming that the former attach to their left while the latter attach to their right, Cameira proposes that the different directions of cliticization of two types of clitic derive from the syntactic structure that is assumed to underlie the relevant constructions. Although both the clitic pronoun and the article are considered to be generated as heads of a Determiner Phrase, when an enclitic pronoun occurs its complement is a *pro*, and thus there is no phonetic item realized to the right of the clitic. This implies that the pronoun must cliticize to the left. By contrast, in the structure with an article the complement is a full Noun Phrase, and consequently there is a lexical item to the right of the clitic. This allows the article to attach phonologically to the right. The two cases are represented in (57), taken from Cameira (1994: 107).



It should be noticed that this analysis may only work if the pronominal clitic is in sentence final position, since if it is followed by a lexical item nothing in principle should prevent the clitic from attaching to the right. In fact, many researchers have argued that phonologically unrealized syntactic

elements cannot affect the application of phonological processes (cf. Nespor and Scorretti 1985; Berendsen 1985; Nespor and Vogel 1986: 2.3.2; Halpern 1991). Moreover, empty elements are usually not included in the range of syntactic information required for the construction of prosodic domains, and, in more elaborated models on the organization of grammar and on the locus of construction of prosodic domains, such as the one assumed in Nespor (1990), the mapping rules for the construction of the prosodic domains higher than the prosodic word are preceded by the deletion of empty elements.⁷³ Thus, under this view, empty elements may not contribute to the construction of prosodic domains, or to the direction of phonological cliticization (see chapter 5 of this book for an alternative proposal to account for the directionality of pronominal clitics).

To support her proposal about the direction of phonological cliticization, Cameira presents the contrasts in (58). The segmental changes in the clitic and the verb illustrated in (58b), which do not occur in the sequence verb-determiner (see 58a), are taken to show the phonological coherence of the sequence verb-pronominal clitic.

- | | |
|---|--|
| (58) a. <i>fiz o palhaço</i>
‘(I) made the clown’
<i>chamar o ladrão</i>
‘(to) call the thief’
<i>pintem o verde</i>
‘paint the green (one)’ | b. <i>fi-lo palhaço</i>
‘(I) made him (a) clown’
<i>chamá-lo ladrão</i>
‘(to) call him (a) thief’
<i>pintem-no verde</i>
‘paint it green’ |
|---|--|

Even though we believe the contrast in terms of direction of cliticization is correct, as we will see in chapter 5, these facts should not be considered purely phonological. In chapter 4, we show that both the form of the enclitic pronoun and the deletion of the verb final consonant cannot be derived by general segmental processes operating at this stage of the language. Instead, we propose that they must be stored in the lexicon as allomorphs and as precompiled phrasal rules, specific to the sequences verb plus pronominal clitic.

In Barbosa (1996), the same kind of data is used to support a specific direction of cliticization of pronominal clitics. In this case, what these facts are intended to show is that preverbal pronominal clitics are not enclitic to a preceding word, since they do not trigger the changes in the form of the clitic and of the preceding word that occur when they are enclitic to the verb. Barbosa therefore claims that preverbal pronouns are phonologically

proclitic to the following verb. In chapter 5, we support this hypothesis resorting to pure phonological processes that are active at this stage of EP.

With respect to the direction of cliticization of pronouns in mesoclitic structures, we have seen in the preceding section that Van der Leeuw (1997) assumes that pronouns are enclitic to the first prosodic word of this construction. In order to explain why the internal clitic is phonologically enclitic rather than proclitic to a prosodic word, Van der Leeuw (1997: 164) adopts Brandão de Carvalho's (1989) proposal that pronominal clitics are always enclitic, whether preverbal or postverbal. As said above, we argue against the latter position in chapter 5, since several facts point to the proclitic status of preverbal pronominal clitics in EP. By contrast, in the cases of mesoclitisis we will empirically support the enclitic status of stressless pronouns (see chapter 6, section 3.6).

Besides the direction of cliticization, a few remarks may also be found in the literature concerning the lexical/postlexical nature of cliticization in EP.

Zwicky (1987) and Halpern (1995) argue that pronominal clitics in EP are in fact inflectional affixes and attach to the verb within the lexical component. This is seen to be the reason why these elements display some affix-like behavior, such as the possibility of appearing before what they consider to be other inflectional affixes in mesoclitic constructions. A review of their arguments is presented in chapter 4, where we further argue against the affixal status of pronominal clitics in this language.

Van der Leeuw (1997) also assumes inflection and cliticization to be similar operations, since they are both considered to be instances of affixation. However, contrary to Zwicky (1987) and Halpern (1995), for Van der Leeuw both inflection and cliticization are postlexical operations, along the lines of Anderson (1992). In opposition to this view, we try to show that inflection and cliticization clearly display different behaviors in EP, and that this can be captured by an analysis whereby the two types of operation apply at different points in the grammar (see chapter 4 of this book, and Vigário 1998c, 1999a, 1999c).

For Crysmann (1999, 2000a, 2000b), pronominal clitics in EP represent a transitional type between morphology and syntax, as they have already acquired the morphological properties of lexical affixes, but, as pointed out in Spencer (1991), their placement is governed by syntax. Among the lexical properties referred in Crysmann's work are: arbitrary gaps, since bound pronominals do not attach to participles; morphological idiosyncrasies, like those described, for example, in Cameira (1994) (see above); and mesoclitisis, which is assumed to be obtained through the morphological opera-

tion of inflection. We have previously addressed these issues as well (Vigário 1998c, 1999a, 1999c). Despite noticing the existence of evidence suggesting the affixal status of clitic pronouns, we have shown that the facts do not necessarily indicate that weak pronouns in EP are derived in the lexical component, on par with other inflectional affixes. Instead, as pointed out above, we believe that there is good evidence that weak pronouns in EP are postlexically attached to their host, contrary to inflectional affixes, which are attached to their morphological base at the lexical component. This topic is the subject matter of chapter 4.

Finally, in d'Andrade and Laks (1992) and Mateus and d'Andrade (2000), cliticization is claimed to be “post-accentual”, since pronominal clitics may originate sequences where word stress is further leftward than the third syllable. This issue is, nevertheless, not developed in these works. In chapter 4, we argue that pronominal clitics are inserted in the string postlexically, and consequently, since word stress is shown to apply within the lexical component, these elements never interact with word stress.

4.3. *Phonological processes and the organization of grammar*

Studies that make explicit assumptions about the organization of the grammar from the perspective of EP phonology include Mateus (1997a), Vigário (1998b, 1999a), Pereira (1999), and, particularly, Mateus and d'Andrade (2000).⁷⁴

Adopting some of the basic claims of lexical phonology (e.g. Kiparsky 1982; Mohanan 1986), Mateus and d'Andrade (2000) argue for the existence of two distinct levels relevant for phonology: lexical and postlexical. Unlike in the classical model, however, the lexicon is not stratified. Instead, the order in word-formation follows from selection restrictions of affixes, which, according to Villalva (1994), may attach to stems, themes or full (inflected) words.

Both the lexical and the postlexical levels may be the locus of morphological operations: derivation and inflection are assigned to the lexicon, while the special case of derivation involving *-mente* and *z*-evaluative suffixes, which attach to words, is argued to belong to the postlexical component.⁷⁵ This division is consistent with EP phonological facts, in that inflected and derived words form single domains for lexical phonology, while words formed with *-mente* and *z*-evaluative suffixes set up two domains for lexical phonological processes.

Phonological rules may also apply either at the lexical or at the postlexical level. They apply within the lexicon if they are conditioned by morphology or depend on idiosyncratic properties of lexical items. This is the case of processes such as theme vowel deletion, vowel harmony, spirantization and velar softening rules. Conversely, phonological rules belong to the postlexical level if they are automatic and free of exceptions. Examples of processes that are considered to be postlexical are resyllabification, the specification of coda segments, default rules to fill underspecified segments and the dissimilation of mid non-back vowels when followed by a palatal segment.

We should remark that EP lexical rules are taken to be structure preserving (cf. Kiparsky 1982; Mohanan 1986), for instance, they are argued not to change one segment into another that does not belong to the underlying system of the language (Mateus and d'Andrade 2000: 105). However, the assignment to the lexical level of processes such as vowel reduction and nasal segments resolution contradicts this assumption, given that these processes create segments that are not present in the underlying system of EP – nasal vowels, schwa and [ɐ] (see chapter 1, section 3.1 for the discussion of structure preservation as a property of lexical processes).

The distinction between lexical and postlexical processes assumed in Mateus and d'Andrade (2000) is also adopted in this book. In chapter 3, we review some of the phonological processes that these authors explicitly assign to one level or another.

5. Concluding remarks

As we have seen, the phonological notion of *word* is absent in much of the work on the phonology of EP. One of our major goals is, therefore, to show that this prosodic domain plays an important role in EP phonology, and that only by considering this constituent may we establish important generalizations in terms of the domain of occurrence of quite a number of processes of the language.

In the realm of the studies on words with two *primary* stresses, the phonological facts brought into the discussion are limited to stress assignment and to the absence of vowel reduction in the first stressed vowel. Nevertheless, other phonological properties seem to characterize these constructions, as we will see in chapter 6. In addition, apart from the recent study of Pereira (1999), the analyses cover just a subset of the existing constructions with two primary stresses. Indeed, derived words

with stressed prefixes, root-compounds, as well as abbreviations, usually fall out of the scope of the constructions that are observed in the literature.

As for clitic words, we could find no systematic characterization of the phonological behavior of clitics. Although examples of stressless words are often provided, an exhaustive definition of the class of phonological clitics is not given. In addition, no compelling arguments for the direction of cliticization or the prosodization of clitics are offered. In fact, in the work where the prosodization of clitics in EP is, at least, mentioned, the great majority of phonological facts that set clitics apart from other words are not given in order to support the actual view that is adopted.

Conversely, many facts that isolate clitics from a phonological point of view are not related in the literature to their phonological status or their particular prosodization. This is clear, for instance, in d'Andrade and Viana's (1993) description of the alternation [j]/0 found in word final position, which is argued not to find an explanation on phonetic or phonological grounds (d'Andrade and Viana 1993: 61).

Considering forms like those in (59), d'Andrade and Viana suggest that, in order to account for the realization of the vowel *e* as [j] in the examples in (59b), it is necessary to refer to the morphological category of pronouns. In fact, as shown by the examples in (59a), *e* in word final position is, in general, absent when followed by a word starting with another vowel.

- | | |
|--|--|
| <p>(59) a. <i>trist(e) amigo</i>
 'sad friend'
 <i>diss(e) o Zé</i>
 said the Zé
 'Zé said'
 <i>viu-lh(e) a careca</i>
 saw him the bold
 'he saw his bold head'</p> | <p>b. <i>disse-o</i>
 '(he) said it'
 <i>o facto de o ver</i>
 the fact that him (to) see
 'the fact that (I) see him'
 <i>deixe-a</i>
 'leave it'</p> |
|--|--|
- (60) *sempr(e) o encontraste*
 always him (you) found
 'you did find him after all'
penso que o president(e) o vetou
 (I) think that the president it vetoed
 'I think the president has vetoed it'

62 *Previous Studies*

Still, in sequences like those in (60), where the second vowel belongs to a pronoun, the non-back vowel is obligatorily deleted, in the same way as in the forms presented in (59a), against the generalization proposed.

D'Andrade and Viana further observe that deletion is not obligatory when the segment that surfaces as a glide belongs to a grammatical word such as *que* or *se*, as illustrated in (61). However, this is in fact true for any (pro)clitic function word, as exemplified in (62), which includes a pronoun and a preposition.

- (61) *se o dia*
if the day
que o Zé
that the Zé

- (62) *já te ouviram*
(they) already you heard
'they already heard you'
de autores
'of authors'

These facts suggest that the morphological category of words is not crucially involved in the (non-)application of vowel deletion.⁷⁶

In contrast with this view, we argue in chapters 3, 5, and 7 that these facts can be accounted for if we acknowledge the prosodic word as the domain of application of final non-back vowel deletion, as well as if we take into consideration the specificity of clitics' prosodization and of the reduction processes that affect clitic words.

As a general conclusion, the prosodic word domain is almost totally absent from the phonological descriptions of EP.⁷⁷ In the following chapters, we hope to show that, besides non-back vowel deletion, many other phonological phenomena crucially refer to this prosodic domain in EP.

⁵⁸ In the examples of this chapter, as well as elsewhere in this book, a vowel bearing word-level stress is indicated with capital letters.

⁵⁹ For the distinction between the secondary stress that has the same general properties as the primary word stress (which is the one we are describing here) and the secondary stress that has rhythmic properties, see, among others, d'Andrade (1997) and Pereira (1999: chap.5), and references therein.

⁶⁰ For a number of phonological similarities and differences between the two varieties of Portuguese, see, for example, Parkinson (1988), Mateus and d'Andrade (2000) and Frota and Vigário (2000). Traditional grammars of Portuguese, such as Said Ali (1964) or Cunha and Cintra (1984), are also rich in references to some of the most salient differences between the two varieties.

⁶¹ Despite the similarities among clitics also found in the European variety of Portuguese, we present in section 5.3.3 a number of facts that show that enclitics and proclitics behave differently from a phonological point of view with respect to their host prosodic word.

⁶² As we will see in section 2.3.2. below, not all linguists accept that preverbal pronominal clitics are phonologically proclitic in EP.

⁶³ On the status of boundaries in SPE framework, see Chomsky and Halle (1968: chap.6) (see also section 1.1, note 1), where it is assumed that the boundaries signaled with = are stronger than formative boundaries (+), but weaker than word boundaries (#). For subsequent arguments against this approach to general juncture phenomena see, in particular, Selkirk (1980: section 3), and references therein, and for specific claims against the boundary =, see Siegel (1980).

⁶⁴ In d'Andrade (1977), only the verbal system is considered in the discussion of word stress, although some remarks can be found on word stress in the nominal system (e.g. d'Andrade 1977: 25).

⁶⁵ This is true at least of *z*-suffixes. In the case of *-mente*, Villalva (1992) assumes that the base this element attaches to is an adjective inflected in the feminine form that establishes a relation of agreement in gender with *mente* (which is diachronically related to the Latin word *mens*, *mentis* 'mind', from which the Portuguese feminine noun *mente* 'mind' is also derived). Villalva therefore proposes that this element is still analyzed as a word that belongs to the category N. We should add, nonetheless, that this analysis is not followed in latter work (Villalva 1994). The status of *z*-evaluative suffixes is not so clear, since, although Villalva considers these suffixes to be *words*, no category label is assigned to these elements, which form a left-headed root compound (see also Villalva 1994 for a different approach).

⁶⁶ Notice that, by d'Andrade's convention (16), if two extrametrical elements exist in the same domain, only the second one remains extrametrical. Thus, unless different clitics are assumed to belong to different domains, in forms such as *falAr-se-te-Á* 'it will be told to you', the first clitic is not extrametrical and therefore the

stress in *lar* may not be derived by the convention (48), which stipulates that the final vowel of the derivational stem receives stress when followed by an extrametrical vowel.

⁶⁷ The following quotation synthesizes Van der Leeuw's reasoning: "...since Optimality Theory abandons rules, repair strategies or derivations, the complete phonological and prosodic structure of all candidates must be posited in one go. This means that there may be candidates with an extra primary stress on the stem and that we may freely insert prosodic boundaries. ...We do not have to derive a form with two prosodic words from one prosodic word, but instead we 'only' have to explain why in this case two prosodic words are better than just one" (Van der Leeuw 1997: 157).

⁶⁸ It should be noticed, however, that the present indicative and the imperfect forms of *haver* do not necessarily (fully) coincide with the endings of verbs in future and conditional (see section 4.3.3, for an alternative proposal on the relation between the auxiliary *haver* and the last part of mesoclitic constructions).

⁶⁹ This is also proposed in Vigário (1999a) and in Mateus and d'Andrade (2000).

⁷⁰ For Pereira, the morphological domain for word stress is the "derivational stem" in the nominal system and the "word" in the verbal system. The derivational stem is defined as a constituent of the word that excludes inflectional affixes. This constituent *may* include one or more roots, as well as derivational affixes (cf. Pereira 1999: 129). Notice that this sort of definition appears to open the way for elements that do not contain a root to be analyzed as derivational stems.

⁷¹ This proposal is reminiscent to the idea suggested in Mateus and d'Andrade (2000) that the prosodic word status of *-mente* and *z*-evaluative suffixes is responsible for the postlexical location of the word-formation process that involves these suffixes (see section 2.3.3, below). It constitutes, in addition, a generalization of Nespor's (1999b) analysis of word+word compounds, according to which these constructions, which include more than one prosodic word, form a phonological phrase in a number of European languages.

⁷² Facing the problems noticed above, presented before in Vigário (1999a, 2000a), Galves (2000) has recently abandoned this idea.

⁷³ See also Truckenbrodt (1999: 1.3) for a detailed argumentation against the visibility of empty syntactic elements and their projections in the mapping between syntax and phonology.

⁷⁴ Three other studies that bear on issues related to the organization of the grammar in Portuguese are Lee (1995) and Schwindt (2000), on morphology and lexical phonology of the Brazilian variety, and Villalva (1994), on EP, but conducted from a morphological point of view.

⁷⁵ As said in the preceding section, the fact that *-mente* and *z*-evaluative suffixes are independent prosodic words seems to be the crucial reason for these constructions to belong to the postlexical level (see Mateus and d'Andrade 2000: 97).

⁷⁶ D'Andrade and Viana note in addition that the glide is obligatory with some numerals. In chapter 6, we show that this is a more general behavior as well, which, like the preceding cases, is also explained by the prosodic configuration that the relevant constructions display. Non-back vowel deletion is particularly difficult to grasp because it interacts with different types of phenomena, as we will see later in this book: the prosodization of clitics and of compound-like sequences (see sections 5.3.3 and 6.2); the reduction of clitic function words (see section 5.3.1.2 and chapter 7); and the lexicalization of combinations of function words (see section 7.4.4.2).

⁷⁷ By adopting Vigário's (1998b, 1999a) analysis of non-back vowel deletion as a prosodic word limit process, Mateus and d'Andrade (2000) constitutes an exception to the generalized absence of the prosodic word domain from EP phonological descriptions. Frota (2000) includes this constituent in the formalization of a number of rules that apply within the intonational phrase domain. However, the prosodic word is not the topic under research. This contrasts with the other prosodic domains, which have been subject to some analysis. For example, (i) in d'Andrade and Viana (1994), Vigário and Falé (1994), Mateus (1995), Freitas (1997), Mateus and d'Andrade (1998, 2000), we can find descriptions of EP syllables, and in most cases proposals concerning syllable structure and the principles underlying its construction; (ii) d'Andrade and Viana (1989), Brandão de Carvalho (1989), Pereira (1990, 1999), and d'Andrade and Laks (1992) refer to the construction of (a constituent similar to) the foot; (iii) Frota (1996, 2000) studies the phonological phrase and the intonational phrase, observing segmental, durational, intonational, and rhythmic phenomena; Falé (1995) and Vigário (1998a) also resort to the phonological phrase and to the intonational phrase in their account of intonational phenomena in this language.

Chapter 3

Phonological phenomena: Description and typology

1. Introduction

In this chapter we survey a number of phonological phenomena of EP that are relevant to the topics under discussion in this book. Our goal is to present a brief description of each phenomenon and to assess its lexical/postlexical nature, as well as the role played by the prosodic word domain in its occurrence.

As pointed out in chapter 2, no distinction is commonly made in EP phonological studies between the prosodic word and the morphosyntactic word. In addition, only in very recent work, such as Mateus (1997a), Vigário (1998b, 1999a), and Mateus and d'Andrade (2000), is there a systematic separation between lexical and postlexical processes. Nevertheless, most of the phonological phenomena that we present in this chapter have long been given some attention by researchers working on EP phonology. In this chapter we present a selection of their work, and in some cases, we propose alternative analyses that we believe to add to the understanding of the relevant phenomena.

Since our primary concern is to identify the locus of operation of phonological phenomena within the grammar, as well as to evaluate the role of the prosodic word domain in their application, no special attention is given to aspects related to the best formalism that could capture the nature of the phenomena analyzed. We will use, in general, SPE-like rule formalism as a descriptive system, interspersed with other systems of representation, such as autosegmental representations, depending either on previous proposals found in the literature, or on the formalism that we may find more transparent for describing a given phenomenon.

Before we proceed, we should make a note on our basic assumptions regarding the underlying system of EP.

In this book, we assume the phonological system of EP shown in (63).⁷⁸ This segmental inventory is in general consistent with the one adopted in Mateus and d'Andrade (2000), the major difference being that we assume the existence of underlying glides, while for Mateus and d'Andrade glides

64 *Phonological phenomena*

always correspond to underlying vowels. We motivate the distinction between underlying and derived glides in section 11.

(63)	Vowels	front	central	back
	high	i		u
	mid	e		o
	low	ɛ	a	ɔ
	Consonants and glides	labial	alveolar	(pre)palatal velar
	glide			j w
	oral stop	p b	t d	k g
	fricative	f v	s z	ʃ ʒ
	nasal	m	n	ɲ
	lateral		l	ʎ
	tap		r	

The same segmental inventory is also assumed in Mateus (1975) and d'Andrade (1977), to the exception of the uvular realization of the trill. In these works, [ʀ] is an underlying segment, while we assume, with Mateus and d'Andrade (2000), that this segment is derived from an underlying /r/ (see section 9).

Also like Mateus (1975), d'Andrade (1977), Andrade (1980), and subsequent work by these authors, we assume that [i] and [ɐ], as well as nasal vowels, do not belong to the underlying segmental inventory of the language and are obtained, instead, through the application of phonological processes such as vowel reduction (cf. section 3), /e/-centralization (cf. section 6), and low vowel raising in nasal environments (cf., for example, Mateus 1975).

2. Word stress

EP word stress assignment has attracted the attention of many researchers, including, among others, Mateus (1975, 1983), d'Andrade (1988),

d'Andrade and Viana (1989), Brandão de Carvalho (1989), Pereira (1990, 1999), d'Andrade and Laks (1992), and Mateus and d'Andrade (2000: chap.6).

Instead of presenting any of the proposals made in the literature, our goal in this section is merely to show some of the facts that have led to the view that word stress assignment refers to the morphological structure of words (as first suggested in Mateus 1983, and developed in subsequent work by d'Andrade and colleagues, and Pereira, cf. chapter 2). In addition, we review some of the data that show the relevance of lexical information for deriving the correct stress placement. This will allow us to support the claim that word stress is assigned in the lexical component (cf. Mateus 1997a; Vigário 1999a; Pereira 1999; Mateus and d'Andrade 2000).⁷⁹

EP word stress can fall on one of the last three syllables of a word, as illustrated in (64).⁸⁰

- | | | | |
|------|--------------------|----------------|------------|
| (64) | a. FINAL | <i>jacarÉ</i> | 'aligator' |
| | b. PENULTIMATE | <i>patAda</i> | 'kick' |
| | c. ANTEPENULTIMATE | <i>pÊssegO</i> | 'peach' |

The forms in (64) show in addition that, within a three syllable window, word stress placement is not predictable on the basis of phonological information, since these words have the same syllabic structure.⁸¹

Although all the words in (64) are nouns, there is a difference in morphological structure between the words in (64a) and those in (64b–c): the form in (64a) shows no overt morphological class-marker, whereas the forms in (64b–c) end in the class-markers *–a* and *–o*, respectively (e.g. d'Andrade 1977; Villalva 1994; Mateus and d'Andrade 2000). This difference has been considered crucial to account for the distinct stress patterns illustrated in (64). Given the generalization that stress is assigned to the last vowel of a nominal stem (e.g. Mateus 1983), stress is expected to be final in nouns with no overt class-marker, and prefinal in nouns with an overt class-marker. This entails that forms with antepenultimate stress are exceptional, and thus the final element of the relevant stems must be lexically marked in order not to receive word stress. However, this is only true for the noun system, as the element that is lexically marked not to bear word stress may become stressed if the relevant word belongs to the verbal category, as the contrasts in (65) show.

- | | | | | |
|------|------------------------|-----------|------------------------|---------------|
| (65) | [dúvida] _N | 'doubt' | [duvIda] _V | '(he) doubts' |
| | [fábrica] _N | 'factory' | [fabrIca] _V | '(he) makes' |

66 *Phonological phenomena*

Without entering into the complex details of the proposals that have been put forward to account for stress assignment in the verbal system, a simple description of the data is suggestive of the impossibility of explaining stress location in verbs strictly on the basis of phonological information.

In the present tenses, stress always falls on the penultimate syllable of the verb form, which may correspond to the last vowel of the verbal root (see 66a), to the theme vowel (see 66b), or to the mood marker (see 66c).⁸²

- (66) a. fal a + o *fAlo* '(I) speak'
 fal a + s *fAlas* '(you) speak'
 fal a *fAla* '(he) speaks'
 fal a + e *fAle* '(he) speak-SUBJ'
 fal a + e + s *fAles* '(you) speak-SUBJ'
 b. fal a + mos *fAlAmos* '(we) speak'
 c. fal a + e + mos *fAlEmos* '(we) speak-SUBJ'

In the past tenses, stress always falls on the theme vowel, regardless of phonological considerations, such as the number or type of syllables after the stress (see 67).

(67)	PERFECT		INDICATIVE IMPERFECT	
2Sg	fal a + ste	falAste	fal a + va + s	falAva
1Pl	fal a + mos	falÁmos	fal a + va + mos	falÁvamos
<hr/>				
	PLUPERFECT		SUBJUNCTIVE IMPERFECT	
2Sg	fal a + ra + s	falAras	fal a + sse + s	falAsses
1Pl	fal a + ra + mos	falÁramos	fal a + sse + mos	falÁssemos

In the future and conditional forms, stress location is also fixed, as it always falls on the vowel following [r], which corresponds historically to the infinitive marker. The examples in (68) illustrate this and further show that stress position cannot be explained by phonological properties of the verb forms.

All these facts indicate that, in order to account for word stress location in EP, it is necessary to have access to morphological and lexical information. As already pointed out, this has led to the proposal that word stress is

assigned within the lexical component, an assumption we also adopt in this book.

(68)	FUTURE	CONDITIONAL
3Sg	falarÁ	falarIa
2Sg	falarÁs	falarIas
1Pl	falarEmos	falarÍamos

As we will see in chapter 5, section 4.3.1 the lexical nature of word stress assignment may explain why pronominal clitics never affect the position of stress in the verbal host, if it is assumed that clitics are combined with their hosts postlexically, as we claim in chapter 4.

Notice, in addition, that the generalization that stress falls within a three syllable stress window, suggested for instance in Van der Hulst (1996) to be universal, can only be maintained if it refers to a lexical domain. In fact, pronominal enclitics behave with respect to prosodic rules as elements incorporated into the verb's prosodic word, as it will be shown in chapter 5. Thus, postlexically, EP allows for prosodic words with stress assigned further leftward than the antepenultimate syllable, as illustrated in (69) (syllable division is marked with dots and is given inside parentheses).⁸³

- (69) *abandonÁramo-la* (a.ban.do.nÁ.ra.mo.la)
 '(we) had abandon her'
participÁvamo-no-la (par.ti.ci.pÁ.va.mo.no.la)
 '(we) have told it to ourselves'

3. Vowel reduction

Detailed phonological analyses of vowel reduction can be found, for example, in Mateus (1975), Andrade (1980), and Brandão de Carvalho (1994) (see also Gonçalves Viana 1883; Lüdtke 1953; and Morais Barbosa 1965, for earlier descriptions of the realization of stressless vowels in EP).⁸⁴

Vowel reduction (VR) is a rather general process that affects stressless vowels, and that consists of (i) the raising and centralization of non-high non-back vowels, (ii) the raising in one degree of the low central vowel, and (iii) the raising of non-high round vowels, as shown in (70).

68 *Phonological phenomena*

(70) Vowel Reduction

$$\begin{array}{ccc}
 /e, \varepsilon/ & /a/ & /o, \text{ɔ}/ \\
 | & | & | \\
 [i] & [ɐ] & [u]
 \end{array}$$

Mateus (1975: 32) formalizes this process as in (71): non-high stressless vowels become non-low and back; and, if they have the same value for back and round, they are further raised to high.

$$(71) \left(\begin{array}{c} V \\ -\text{high} \\ \langle \alpha\text{back} \rangle \\ \alpha\text{round} \\ -\text{stress} \end{array} \right) \rightarrow \left(\begin{array}{c} < +\text{high} > \\ -\text{low} \\ +\text{back} \end{array} \right)$$

VR is illustrated in the forms in (72), taken from Mateus (1975: 28–29), where the same underlying vowel appears in stressed and unstressed position.

(72)	<i>dEdo</i>	[e]	<i>dedAda</i>	[i]	<i>batEr</i>	[e]	<i>bAte</i>	[i]
	‘finger’		‘pinch’		‘(to) hit’		‘(he) hits’	
	<i>fEstá</i>	[ɛ]	<i>festInha</i>	[i]	<i>mEl</i>	[ɛ]	<i>melAdo</i>	[i]
	‘caress’		‘caress-DIM’		‘honey’		‘honey-colored’	
	<i>gAto</i>	[a]	<i>gatInho</i>	[ɐ]	<i>cAla</i>	[a]	<i>calAra</i>	[ɐ]
	‘cat’		‘cat-DIM’		‘(he) silences’		‘(he) had silenced’	
	<i>fOgo</i>	[o]	<i>foguEira</i>	[u]	<i>mOvo</i>	[o]	<i>movEr</i>	[u]
	‘fire’		‘bonfire’		‘(I) move’		‘(to) move’	
	<i>lOja</i>	[ɔ]	<i>lojIsta</i>	[u]	<i>fOco</i>	[ɔ]	<i>focAdo</i>	[u]
	‘shop’		‘shopkeeper’		‘focus’		‘focused’	

There are many exceptions to VR, as shown in (73). The examples in (74) show, furthermore, that the same underlying vowel may appear reduced or unreduced in morphologically and semantically related words. This suggests that the non-application of VR has to be lexically marked for specific words.⁸⁵

- | | | | | | | |
|------|-------------------|-----|--------------------|-----|-------------------|-----|
| (73) | <i>relatOr</i> | [a] | <i>direccionAl</i> | [ε] | <i>persuasIvo</i> | [a] |
| | ‘reporter’ | | ‘directional’ | | ‘persuasive’ | |
| | <i>metrificAr</i> | [ε] | <i>procuraçãO</i> | [ɔ] | <i>redacçãO</i> | [a] |
| | ‘metrify’ | | ‘procuration’ | | ‘writing’ | |
| | <i>mestrAdo</i> | [ε] | <i>retÓrica</i> | [ε] | <i>motInha</i> | [ɔ] |
| | ‘masters’ | | ‘rhetoric’ | | ‘motorcycle-DIM’ | |
| | | | | | | |
| (74) | <i>economIa</i> | [ɔ] | <i>econÓmico</i> | [u] | | |
| | ‘economy’ | | ‘cheap’ | | | |
| | <i>velharIa</i> | [ε] | <i>velhIce</i> | [i] | | |
| | ‘old stuff’ | | ‘ageness’ | | | |
| | <i>armazÊm</i> | [a] | <i>armazenAr</i> | [ø] | | |
| | ‘warehouse’ | | ‘(to) store up’ | | | |
| | <i>mestrAdo</i> | [ε] | <i>mestrIa</i> | [i] | | |
| | ‘masters’ | | ‘excellence’ | | | |
| | <i>festInha</i> | [ε] | <i>festejAr</i> | [i] | | |
| | ‘party-DIM’ | | ‘(to) celebrate’ | | | |

The pairs of words in (75), which are phonologically identical but have a different meaning, behave differently with respect to VR. This indicates that the absence of vowel reduction in these cases is not explained by the phonological properties of words.

- | | | | | |
|------|--------------|-----|-----------------|----------|
| (75) | <i>fEstá</i> | [ε] | <i>festInha</i> | [i]/[ε] |
| | ‘caress’ | | ‘caress- DIM’ | |
| | <i>fEstá</i> | [ε] | <i>festInha</i> | *[i]/[ε] |
| | ‘party’ | | ‘party- DIM’ | |

Besides the vowels included in (70) above, it is often observed that /i/ is also centralized to schwa and thus is also subject to VR. However, this phenomenon is distinct from the one that affects other non-back vowels because it only occurs regularly in word final position. Additionally, /i/-centralization only affects vowels that are *verb* final.⁸⁶ This is illustrated in the forms in (76a), which contrast with the nominal forms presented in (76b).⁸⁷ It should be noticed that the theme vowel /i/ in verb final position is also subject to a special process of deletion, when preceded by /z/, as shown in (77).

70 *Phonological phenomena*

- (76) a. (*partIr* *pArte* [i] b. *jÚri* [i] *GÂndi* [i]
 ‘(to) leave’ ‘(he) leaves’ ‘jury’ ‘Gandhi’
 (*dormIr*) *dOrme* [i] *tÁxi* [i] *DÍli* [i]
 ‘(to) sleep’ ‘(he) sleeps’ ‘taxi’ ‘id.’
 (*abrIr*) *Abre* [i] *BÂmbi* [i] *pÉni* [i]
 ‘(to) open’ ‘(he) opens’ ‘Bambi’ ‘penny’
 (*cobrIr*) *cObre* [i] *dÉrbi* [i] *dÂndi* [i]
 ‘(to) cover’ ‘(he) covers’ ‘special match’ ‘dandy’
- (77) (*reproduzIr* ‘(to) reproduce’) *reproduz* ‘(he) reproduces’
 (*reluzIr* ‘(to) shine’) *reluz* ‘(it) shines’
 (*conduzIr* ‘(to) drive’) *conduz* ‘(he) drives’

We therefore take /i/-centralization to be specific to verbs and thus keep the characterization of VR as a process applying only to non-high vowels.

Despite the existence of many unruléd exceptions to VR, there are also a number of contexts where the non-application of VR is predictable. These are related to syllable structure, to the position of the vowel/syllable within the word and to morphological information. Specifically, VR does not apply:

(i) to vowels in a branching nucleus (see 78);⁸⁸

- (78) *auditorIa* [a] *sinusoidAl* [ɔ]
 ‘assessment’ ‘id.’
 oitAvo [o] *pneumÁtico* [e]
 ‘eighth’ ‘pneumatic’

(ii) to nasal vowels (see 79);

- (79) *conduzIr* [õ] *alongAr* [õ]
 ‘(to) drive’ ‘(to) lengthen’
 quentInho [ẽ] *amendoAdo* [ẽ]
 ‘warm-DIM’ ‘oval’

(iii) to non-round vowels in syllables closed by /l/ (see 80);⁸⁹

- (80) *delgAdo* [ɛ] *crueldAde* [ɛ]
 ‘thin’ ‘cruelty’
balldo [a] *lealdAde* [a]
 ‘abandoned’ ‘loyalty’

(iv) to non-central word initial vowels (see 81);⁹⁰

- (81) *erguer* [e]/[i] *odiAr* [o]/[ɔ]
 ‘(to)raise’ ‘(to) hate’
eminEnte [e]/[i] *omitIr* [o]/[ɔ]
 ‘close’ ‘(to) silence’

(v) to vowels in word final syllables closed by a sonorant consonant (see 82);⁹¹

- (82) *Álcool* [ɔ] *líder* [ɛ] *Éden* [ɛ]
 ‘alcohol’ ‘leader’ ‘paradise’
incrível [ɛ] *repórter* [ɛ] *pólen* [ɛ]
 ‘incredible’ ‘reporter’ ‘pollen’

(vi) to the stem final vowel of stem compounds (see 83);⁹²

- (83) *cIne-radiografla* [ɛ] *tÉle-chamAda* [ɛ]
 ‘X-ray movie’ ‘distant call’
Auto-serviço [ɔ] *vÍdeo-conferÊncia* [ɔ]
 ‘car service’ ‘video conference’

(vii) to vowels that end clipped words or that belong to loanwords that have not yet been completely integrated in the language system (see 84).⁹³

- (84) *Expo* [ɔ] *Euro* [ɔ]
 ‘exposition’ ‘euro; european cup’
mIcro [ɔ] *mOto* [ɔ]
 ‘microphone, micro-waves’ ‘motorcycle’

Except for the (regular) cases listed above, VR seems to apply without true exceptions only in post-tonic position (see also Marquilhas 2000).

The lexical nature of VR is proposed in Vigário (1999a) and Mateus and d'Andrade (2000) on the basis of the existence of a large number of (unpredictable) exceptions to its application.⁹⁴ In addition to this, the fact that VR is conditioned by morphological information, as in the cases illustrated in (83) above, also points to its lexical nature.

Another piece of evidence suggesting the lexical nature of VR comes from the fact that it applies also to vowels in syllables that become closed due to (optional) schwa deletion, as illustrated in (85).

- (85) *alegria* [ɐˈlɛɡɾiɐ] < /alegria/ 'joy'
 valetinha [vɐˈlɛtɨɲɐ] < /valetɨna/ 'drain-DIM'

In these examples, the underlying /ɛ/ and /e/, which can also surface as a schwa, are deleted due to their unstressed status. The syllable structure of the words becomes *al.gri.a*, and *val.ti.nha*, as supported by the fact that coda *l*-velarization applies. Despite the fact that the initial syllable of these words has become closed, VR is not blocked (see also Andrade 1980: 4.2.1). This behavior may be accounted for if we assume that VR is a lexical operation, while optional schwa deletion, syllable restructuring, and *l*-velarization are postlexical phenomena.

VR is not affected by the presence of word initial non-primary stress or emphatic stress either. These will be argued in sections 16 and 17 to be postlexically assigned. Thus, the assumption that VR is turned off postlexically explains why vowels bearing postlexically assigned stress undergo, nevertheless, vowel reduction.

Finally, VR must precede postlexical resyllabification, since word final sonorants may surface at the onset of a following syllable starting with a vowel and, while at this level the context for the application of VR obtains, vowels in such environment never undergo VR (see 86).

- (86) *líderAnça* (li.de.ran.ça) *[ɛ]/[i] 'leadership'
 líder (li.der) [ɛ]/*[i] 'leader'
 líder adorado (li.de.ra.do.ra.do) [ɛ]/*[i] 'adored leader'

The same behavior is also found in word initial vowels that may become non-initial due to postlexical resyllabification. As mentioned above (see iv), VR usually operates word internally but not word initially. The fact that a stressless vowel becomes non-initial due to postlexical resyllabification – which causes the final consonant of the preceding word to become the

onset of the following word initial empty onset – shows that the process is not active postlexically. This is illustrated in (87).

- | | | | | |
|------|-----------------------|-----------------------|---------------|-----------------------|
| (87) | <i>levantAr</i> | (le.van.tar) | *[ɛ]/*[e]/[i] | ‘raise’ |
| | <i>emotIvo</i> | (e.mo.ti.vo) | [e]/*[i] | ‘emotional’ |
| | <i>animal emotIvo</i> | (a.ni.ma.le.mo.ti.vo) | [e]/*[i] | ‘emotional
animal’ |

In conclusion, the facts observed in this section clearly indicate that vowel reduction is not operative postlexically.

As we will see in chapter 5, section 4.1.2, vowel reduction also applies to some monosyllabic words. In order for VR to operate here, it must be assumed that these words lack word-stress and therefore do not form prosodic words. This implies that VR does not have the prosodic word as its domain of application.

4. Theme vowel deletion

When followed by another vowel, the theme vowel of verbal forms is obligatorily deleted (e.g. Mateus 1975: 2.1; d’Andrade 1977: 4.2, 1994d). This process is commonly accepted to apply in the present indicative and subjunctive verb forms, as illustrated in (88).⁹⁵

- | | | | |
|------|-----------------------------|-----------------------------|--------------------------------------|
| (88) | <i>falo</i> (/fal a + u/) | <i>fale</i> (/fal a + e/) | <i>falemos</i> (/fal a + e + mus/) |
| | ‘(I) speak’ | ‘(he) speak-SUBJ’ | ‘(we) speak-SUBJ’ |
| | <i>bato</i> (/bat e + u/) | <i>bata</i> (/bat e + a/) | <i>batamos</i> (/bat e + a + mus/) |
| | ‘(I) hit’ | ‘(he) hit-SUBJ’ | ‘(we) hit-SUBJ’ |
| | <i>parto</i> (/part i + u/) | <i>parta</i> (/part i + a/) | <i>partamos</i> (/part i + a + mus/) |
| | ‘(I) leave’ | ‘(he) leave-SUBJ’ | ‘(we) leave-SUBJ’ |

In EP there is only one additional inflectional morpheme that clearly starts with a vowel and is preceded by the theme vowel, namely, the imperfect suffix of the 2nd and 3rd conjugations /ia/, as illustrated in (89).

- | | | | | |
|------|-----------------|--------------------|-------------------|---------------------|
| (89) | <i>batia</i> | (/bat e + ia/) | <i>partia</i> | (/part i + ia/) |
| | ‘(he) hit-IMP’ | | ‘(he) leave-IMP’ | |
| | <i>batias</i> | (/bat e + ia + s/) | <i>partias</i> | (/part i + ia + s/) |
| | ‘(you) hit-IMP’ | | ‘(you) leave-IMP’ | |

While theme vowel deletion may also be seen to apply in these cases, Mateus and d'Andrade (2000: 77) assume that here the two adjacent non-back vowels undergo a specific derivation: raising of /e/ to [i], in the case of the verbs of the 2nd conjugation, and fusion of the two high vowels, which become a single [i].⁹⁶ Although the deletion of the theme vowel yields a simpler analysis, the fact that stress falls, superficially, on the tense morpheme, would constitute a counter-example to the generalization that, in past tenses, verbal stress does not fall on tense morphemes (see section 2).

Whether or not theme vowel deletion is likely to be generalized to all tense paradigms, it seems clear that it does not apply to sequences of two vowels in non-verbal inflectional environments. This is illustrated in (90), where vowel deletion is not allowed.

(90)	<i>rÓseo</i>	*0	<i>rÓsea</i>	*0	<i>areAl</i>	*0
	'rosy-MASC'		'rosy-FEM'		'sand field'	
	<i>família</i>	*0	<i>esvoaçAr</i>	*0	<i>moAgem</i>	*0
	'family'		'(to) flutter'		'milling'	

The process is formalized in Mateus et al. (1990: 383) as in (91).⁹⁷

$$(91) \quad V \rightarrow \emptyset \quad / \quad \left. \begin{array}{c} - \\ \text{Theme} \end{array} \right\} + V$$

The morphological conditioning of this process argues for its lexical status, as claimed by Mateus and d'Andrade (2000). In addition, the prosodic word plays no role in its application.

The fact that theme vowel deletion requires that the vowel that follows the theme vowel belongs to a verb inflectional affix will be used in chapter 4 as a diagnostic for the affix/clitic status of stressless pronouns in EP (see in particular section 3.2).

5. Final nasal diphthongization

Nasal consonants in general do not appear in coda position in EP. The way underlying nasal segments in syllable final position are parsed in this language is a rather complex one, where both phonological and morphological information play a role.

The complexity of the phenomenon has attracted the attention of many researchers, such as Morais Barbosa (1965), Mateus (1975), Brandão de

Carvalho (1988), d'Andrade and Kihm (1988), Morales-Front and Holt (1997), Mateus and d'Andrade (2000), among others. Our goal in the present section is to describe some of the peculiarities of the phenomenon, rather than to propose a full account of it. Nevertheless, the presentation of the facts in the following paragraphs certainly reflects our own understanding of the phenomenon of nasality, which is inspired to a great extent by Mateus (1975), Morales-Front and Holt (1997) and Mateus and d'Andrade (2000).

Along the lines of d'Andrade and Kihm (1988), we assume that surface nasal vowels always derive from a sequence of vowel plus an underlying underspecified nasal segment, which we represent below as [+nas]. In addition, we adopt Morales-Front and Holt's (1997) assumption that the underlying form of the nominal plural suffix is /s/ rather than /es/ (unlike, for example, Mateus 1975) and words like *pão* 'bread' have no underlying final /e/ (unlike d'Andrade 1977: chap.3).

In EP, an underspecified nasal segment in morpheme final position behaves in different ways depending on morphological and phonological conditions:

- (i) if the nasal segment is followed by a sonorant consonant and is part of the prefix *in-*, the nasal consonant is deleted (cf. 92.1a);
- (ii) if the nasal segment is followed by a nasal consonant and is part of the prefix *en-*, the nasal segment is deleted (cf. 92.1b);⁹⁸
- (iii) if the nasal segment belongs to a prefix such as *pan-*, and is followed by a(ny) consonant, the preceding vowel is nasalized and the nasal segment is deleted (cf. 92.1c); the same happens in derived environments with the prefix *in-* when followed by non-sonorant consonants and with *en-* when followed by an oral consonant, as well as in underived environments;
- (iv) if the nasal segment is followed by a vowel:
 - a. in derived environments, the nasal segment surfaces as an alveolar consonant in the onset of the following syllable and the preceding vowel is oral (cf. 92.2a);
 - b. in inflectional environments, the nasal segment nasalizes the preceding vowel and is deleted (cf. 92.2b);
- (v) if the nasal segment is followed by a consonant in final position, necessarily the fricative that marks the plural in the nominal system, the nasal segment surfaces as a glide with the same value for the feature [back] as the following fricative (cf. 92.3);⁹⁹ the glide is then

76 *Phonological phenomena*

associated with the preceding nucleus and the nucleus vowel becomes nasalized.

- (vi) if the nasal segment is in word final position, it surfaces as a nasal glide with the same value for the feature [back] as the preceding vowel and nasality spreads to the syllabic nucleus, as illustrated in (92.4).¹⁰⁰

(92) 1. nasal segment followed by another consonant

- (a) prefix *in-* + sonorant:

<i>irreal</i>	[i]	(/i[+nas]+.../)	‘unreal’
<i>illegal</i>	[i]	(/i[+nas]+.../)	‘illegal’

- (b) prefix *en-* + nasal consonant:

<i>enobrecer</i>	[e]	(/e[+nas]+.../)	‘(to) dignify’
------------------	-----	-----------------	----------------

- (c) prefix *pan-* and underived environments:

<i>panromânico</i>	[ẽ]	(/pa[+nas]+.../)	‘cross-romanian’
<i>mendigo</i>	[ẽ]	(/me[+nas]+.../)	‘begger’

2. nasal segment followed by a vowel

- (a) in derived environments:

<i>inutilizar</i>	[n]	(/i[+nas]+.../)	‘nullify’
<i>enaltecer</i>	[n]	(/e[+nas]+.../)	‘praise’
<i>irmanar</i>	[n]	(/irma[+nas]+.../)	‘fraternize’

- (b) in inflectional environments:

<i>irmão</i>	[ẽ]	(<[ã]) (/irma[+nas]+u/)	‘brother (MASC)’
<i>irmãos</i>	[ẽ]	(<[ã]) (/irma[+nas]+u+s/)	‘brother (MASC-PL)’

3. nasal segment followed by a word final consonant

<i>cães</i>	[ẽ]	(<[ã]) (/ka[+nas]+s/)	‘dog (PL)’
<i>bens</i>	[ẽ]	(<[ẽ]) (/be[+nas]+s/)	‘good (PL)’

4. nasal segment in word final position

<i>falām</i>	[ẽw̃]	(<[ãw̃]) (/fal a+[+nas]/)	‘(they) speak’
<i>batem</i>	[ẽ]	(<[ẽ]) (/bat e+[+nas]/)	‘(they) hit’
<i>cão</i>	[ẽw̃]	(<[ãw̃]) (/ka[+nas]/)	‘dog’
<i>bem</i>	[ẽ]	(<[ẽ]) (/be[+nas]/)	‘well’

Regardless of the actual analysis that is proposed to account for nasality resolution in EP or the precise details of its description, it seems clear that this process is morphologically conditioned and it displays idiosyncratic behavior. This has led Mateus and d’Andrade (2000: 7.1) to locate this process at the lexical level.¹⁰¹ By contrast, final nasal diphthongization,

treated in Mateus and d'Andrade (2000: 7.2) as a glide insertion process, is left unclassified.

The analysis adopted here for the forms in (92.4), which is inspired in Mateus (1975), and Mateus and d'Andrade (2000), is supported by one revealing fact that suggests its productive application.¹⁰² In the construction with 'haver de', for some speakers, the preposition *de* has morphologically incorporated into the auxiliary, when the verb form is monosyllabic. As a result, the person/number suffixes are realized after the incorporated preposition, instead of appearing after the verb stem and before the preposition. This is illustrated in (93).

(93)	Standard	Reanalyzed	
a.	<i>hás-de</i>	a'. <i>hades</i>	'(you) will'
b.	<i>hão-de</i> [ẽw̃] (<[ãw̃])	b'. <i>hadem</i> [ẽj] (<[ẽj])	'(they) will'

The forms in (93b–b') show how the value of the [back] feature of the glide is dependent on the value of the [back] feature of the preceding vowel. In fact, these forms are interesting because in each case the glide is preceded by a vowel with distinct values for the [back] feature, which correlate with the [back] feature value of the glide: in the standard form, the final glide is preceded by a back vowel and thus it surfaces as back, whereas in the form where the preposition has incorporated into the verb, the glide is preceded by a non-back vowel and thus it surfaces as non-back.

Although final diphthongization is a quite general phenomenon, there are a few arguments that point to its lexical application.

First, the existence of words such as *irmã* 'sister' or *orfã* 'orphan-FEM', where no final nasal glide occurs, shows that final nasal diphthongization has to refer to morphological structure. Indeed, if the nasal glide was inserted postlexically, these words should surface with a final [w̃]. Under the view assumed here, these forms follow the pattern of other words that have a class-marker, such as *irmão* 'brother': assuming that in these cases the nasal segment is followed by the class-marker *-a*, the nasal segment nasalizes the preceding vowel and is deleted; the two underlying identical vowels (/a/) then fuse and surface as a single nasal vowel (see Mateus et al. 1990 for a similar view). In these cases, therefore, there is no final nasal segment, and therefore no nasal glide surfaces.¹⁰³

Additionally, as we will see in chapter 4, verbs ending in nasal diphthongs are realized in the same way when followed by an enclitic pronoun or another word, as illustrated in (94).

(94) <i>falam-lhes</i>	[ẽw̃] / *[ẽ] / *[e]	‘(they) talk to them’
<i>comem-no</i>	[ẽj] / *[ẽ] / *[e]	‘(they) eat it’
<i>comem ovos</i>	[ẽj] / *[ẽ] / *[e]	‘(they) eat eggs’
<i>comem nozes</i>	[ẽj] / *[ẽ] / *[e]	‘(they) eat nuts’

This again suggests that the relevant domain for nasal diphthongization is the morphological word, and/or that this is a lexical process, since there is evidence that pronominal clitics attach to the verb postlexically (see chapter 4) and that they incorporate into the preceding prosodic word (see chapter 5). By contrast, if this process were considered to apply postlexically at the right-edge of the prosodic word, the verb should not surface with a nasal diphthong because the incorporation of the clitic causes the verb final syllable to become non-final with respect to the prosodic word domain.

Finally, the process also applies to a clitic function word – *em* [ẽj] ‘in’ – thus confirming that it applies with reference to the morphological word rather than the prosodic word.

We therefore conclude that final nasal diphthongization operates within the lexical component and does not refer to the prosodic word.¹⁰⁴

6. /e/-centralization

According to Mateus (1975: 1.3.3) there are two distinct processes of centralization that affect front mid vowels in EP when followed by palatal segments. One of these processes applies when /e/ is followed by a palatal glide, whether in stressed or in stressless positions (see 95).

(95) a. <i>arEia</i>	[ẽj]	‘sand’	b. <i>leitarIa</i>	[ẽj]	‘creamery’
<i>passEio</i>	[ẽj]	‘walk’	<i>despeitAr</i>	[ẽj]	‘(to) vex’

The rule proposed by Mateus is shown in (96).

$$(96) \quad \begin{pmatrix} \text{V} \\ \text{-high} \\ \text{-back} \end{pmatrix} \rightarrow \begin{pmatrix} \text{-syll} \\ \text{-low} \\ \text{+back} \end{pmatrix} \quad / \quad - \begin{pmatrix} \text{-cons} \\ \text{-back} \end{pmatrix}$$

The other process only applies to the front mid vowel in stressed positions, as shown by the contrasts in (97–98), also taken from Mateus (1975: 34).

- (97) *lEnha* [ɐ̃] ‘firewood’ *cerEja* [ɐ̃] ‘cherry’
 tElha [ɐ̃] ‘tile’ *fEcho* [ɐ̃] ‘(I) close’
- (98) *lenhadOr* [i]/*[ɐ̃] ‘woodcutter’ *cerejEira* [i]/*[ɐ̃] ‘cherry tree’
 telhAdo [i]/*[ɐ̃] ‘roof’ *fechadUra* [i]/*[ɐ̃] ‘lock’

The rule Mateus assumes in order to account for these facts is given in (99).

$$(99) \begin{pmatrix} V \\ -\text{high} \\ -\text{back} \\ +\text{stress} \end{pmatrix} \rightarrow \begin{pmatrix} -\text{low} \\ +\text{back} \end{pmatrix} / - \begin{pmatrix} C \\ -\text{ant} \\ -\text{back} \end{pmatrix}$$

Andrade (1994a) merges in a single process the centralization of /e/ in stressed position followed by a palatal glide and the centralization of /e/ in stressed position followed by a palatal consonant. This author also refines Mateus’ formulation of the rule, so that it does not include low vowels.¹⁰⁵ This is a necessary restriction in the dialect we are describing, for /e/ in this dialect does not centralize (cf. *velha* [véʎe]/*[véʎe] ‘old-FEM’; *Beja* [béʎe]/*[béʎe] ‘id.’). D’Andrade’s rule is shown in (100).

$$(100) \begin{pmatrix} V \\ -\text{high} \\ -\text{low} \\ +\text{stress} \end{pmatrix} \rightarrow [+back] / - \begin{pmatrix} +\text{high} \\ -\text{back} \end{pmatrix}$$

Our approach to this phenomenon will make use of both analyses. In fact, although it would be desirable to merge in a single rule the two cases of centralization, we believe that some facts argue for the separation of the two phenomena.

Starting our discussion with d’Andrade’s analysis, we should point out that the formulation in (100) excludes the rule from applying in unstressed environments, although, as noticed by the author, it may also affect unstressed front mid vowels, as also shown by the examples in (95b) above.

Along the lines of Mateus (1975), our understanding of /e/-centralization is that there is a centralization process that is triggered by the presence of a tautosyllabic palatal glide (we will call this process *tautosyllabic /e/-centralization*). Unlike Mateus, however, and following d’Andrade (1994a), we believe that the other process of centralization may be

triggered by any palatal segment, whether a consonant, a (heterosyllabic) glide or a front vowel (we will call this process *heterosyllabic /e/-centralization*).

This approach allows us to straightforwardly explain why syllable final fricatives, which always surface as palatal, never trigger tautosyllabic /e/-centralization, since this phenomenon is seen to apply only when a palatal glide follows the vowel. The absence of tautosyllabic /e/-centralization in fricative environments is illustrated in (101), taken from d'Andrade (1994a).

(101) <i>incEsto</i>	[ɛʃ]/*[ɐʃ]	'incest'	<i>rEsma</i>	[ɛʒ]/*[ɐʒ]	'ream'
<i>vEspa</i>	[ɛʃ]/*[ɐʃ]	'wasp'	<i>mEsma</i>	[ɛʒ]/*[ɐʒ]	'really'
<i>tEsto</i>	[ɛʃ]/*[ɐʃ]	'potlid'	<i>lEsma</i>	[ɛʒ]/*[ɐʒ]	'slow-moving'

So as to account for the non-application of the process in the forms in (101), d'Andrade proposes that centralization applies before the syllable final fricative becomes palatal. The latter state of affairs is the result of the specification of the place features of an underlyingly underspecified coda fricative. However, according to Morales-Front and Holt (1997) the word final fricative must be specified for the place features in order to explain the specific palatal realization of /l/ and of nasal segments in the plural forms of words ending with lateral and nasal segments, respectively (e.g. *animal/animais* 'animal(PL)', *pão/pães* 'bread(PL)' – see also section 5). Since these realizations are morphologically conditioned, they must be assumed to be obtained in the lexical level. Pointing in the same direction is the fact that postlexical resyllabification may affect the coda fricative and remove the context for the feature specification process. Therefore, if this proposal is adopted (as done in this book, as well as in other works, such as Mateus and d'Andrade 2000), d'Andrade's analysis is no longer available. Tautosyllabic /e/-centralization may thus be seen to be triggered by tautosyllabic palatal glides but not tautosyllabic palatal fricatives.

Another reason to separate the two phenomena, is that this centralization process seems to apply without exceptions, unlike the second type of centralization, as we will see below. Notice, furthermore, that in some dialects only tautosyllabic palatal glides, but not heterosyllabic palatal segments, trigger centralization. We therefore assume that tautosyllabic /e/-centralization applies within the syllable domain when /e/ is followed by a palatal glide (see 102).

$$(102) \begin{pmatrix} V \\ -\text{high} \\ -\text{low} \end{pmatrix} \rightarrow [+back] / \left[\dots - \begin{pmatrix} -\text{syll} \\ -\text{cons} \\ -\text{back} \end{pmatrix} \dots \right]_{\sigma}$$

As for the second type of centralization, which is the one that will be most relevant for our discussion in this book, we believe it may also apply to both stressed and unstressed front mid vowels. The asymmetry in the behavior of stressed and unstressed vowels noticed in Mateus (1975) may, in fact, be account for if we take into consideration the application of vowel reduction. Indeed, the cases where /e/-centralization is not observable is when vowel reduction causes the raising and centralization of unstressed front vowels. If vowel reduction applies, therefore, there is simply no front vowel to undergo centralization. Notice, furthermore, that the result of vowel reduction, like /e/-centralization, is also a *central* vowel. Thus, the application of vowel reduction yields an output similar to /e/-centralization in that it also allows for the dissimilation of the sequence of two adjacent underlying palatal segments.

Following d'Andrade's view, we assume that the phenomenon of centralization we are describing is affected by any (heterosyllabic) palatal segment. That all palatal consonants trigger the process was already shown in Mateus (1975) and illustrated in (97) above. The examples in (103) further show that front high vowels also trigger heterosyllabic /e/-centralization.

- (103) *veÍculo* [vɛi] 'vehicle' *europɛizAr* [vɛi] '(to) Europeanize'
cheÍnho [vɛi] 'full-DIM' *reiterAr* [vɛi] '(to) repeat'
deíficAr [vɛi] '(to) deify' *homogeneízAr* [vɛi] 'homogenize'

Thus, we propose that centralization applies when a front mid vowel is followed by a heterosyllabic high palatal segment within the prosodic word domain, as in (104).¹⁰⁶

$$(104) \begin{pmatrix} V \\ -\text{high} \\ -\text{low} \end{pmatrix} \rightarrow [+back] / \left[\dots - \begin{pmatrix} +\text{high} \\ -\text{back} \end{pmatrix} \dots \right]_{\omega}$$

(where the two adjacent segments are heterosyllabic)

82 *Phonological phenomena*

As said above, unlike tautosyllabic /e/-centralization, heterosyllabic /e/-centralization has idiosyncratic exceptions, like those in (105).

- (105) *mExo* [eʃ]/*[ɐʃ] ‘(I) touch’ *mExas* [eʃ]/*[ɐʃ] ‘(you) touch-SUBJ’
remExo [eʃ]/*[ɐʃ] ‘(I) jumble’ *remExa* [eʃ]/*[ɐʃ] ‘(he) jumble-SUBJ’
rEjo [eʒ]/*[ɐʒ] ‘(I) conduct’ *rEja* [eʒ]/*[ɐʒ] ‘(he)conduct-SUBJ’

Notice that we cannot assume that centralization does not operate in these cases because, underlyingly, the stressed vowel is low and the mid realization of stressed vowels in (105) is the result of a specific process of vowel harmony.¹⁰⁷ Indeed, the forms in (106) also undergo vowel harmony, and centralization applies because the vowel is followed by a palatal fricative.

- (106) *protEjo* [ɐ] ‘(I) protect’
elEjo [ɐ] ‘(I) elect’

The fact that this process has exceptions argues for its lexical status.¹⁰⁸

An independent argument for the lexical operation of /e/-centralization is provided by the absence of centralization when a strictly optional [j] is inserted in the string to break a hiatus in nasal environments, as in (107). Notice that in these cases, the glide clearly tends to be syllabified in the onset of the syllable that follows /e/.¹⁰⁹

- (107) *dÊem* [eʃ]/[ejʃ]/*[ɐjʃ] ‘(they) give-SUBJ’
lÊem [eʃ]/[ejʃ]/*[ɐjʃ] ‘(they) read’
vÊem [eʃ]/[ejʃ]/*[ɐjʃ] ‘(they) see’

Assuming that the optional glide is postlexically inserted and that heterosyllabic /e/-centralization is a lexical process accounts for the impossibility of centralization in this context in this variety of EP, despite the fact that the segmental context for the rule is met postlexically.

As we will see in chapter 4, the absence of heterosyllabic /e/-centralization between a verb and an enclitic pronoun, which together form a single prosodic word (as supported by a number of processes described in chapter 5, section 4.3.1), also suggests that the locus of operation of this process is the lexical component.

7. Glide insertion to break a hiatus

The forms in (108) illustrate a general phenomenon of palatal glide insertion between two vowels, where the first one is stressed. In (108a) the glide is not present because the first vowel is stressless, whereas in (108b) the glide is inserted because there are two adjacent vowels and the first one is stressed.

(108) a.	<i>areAl</i>	[já]	b.	<i>arEia</i>	[éjɐ]
	‘sand field’			‘sand’	
	<i>passeAr</i>	[já]		<i>passEio</i>	[éjɐ]
	‘(to) walk’			‘walk’	
	<i>ceAr</i>	[já]		<i>cEia</i>	[éjɐ]
	‘to sup’			‘supper’	

Contrasting forms like those in (108) with those in (109), where no glide is inserted, Mateus (1975: 1.3.5.1) characterizes the first vowel (V1) involved in this process as a stressed front mid vowel and the second vowel (V2) as a non-high back vowel.

(109)	<i>batEu</i>	[éw]	‘(he) hit-PAST’
	<i>ilÍaco</i>	[íɐ]	‘Iliac’

The data we could find are compatible with this characterization of V1. The following forms show the absence of glide insertion when the first vowel is round.

(110)	<i>vOa</i>	[óɐ]/*[ójɐ]	‘(he) flies’	<i>vOo</i>	[óu]/*[óju]	‘fly’
	<i>dOa</i>	[óɐ]/*[ójɐ]	‘(he) donates’	<i>dOo</i>	[óu]/*[óju]	‘(I) donate’

As for V2, it is not clear that it must be non-high. On the one hand, in the form *bateu*, the stressed vowel is not followed by a vowel, but rather by a segment that is obligatorily realized as a glide. On the other hand, the final vowel of words like *passeio* always surfaces as high, and thus it seems at this moment unmotivated to postulate an underlying non-high vowel (see note 95).

It is clear, by contrast, that V2 has to be a vowel, rather than a glide, and, in addition, it can also be a back high vowel (or a front mid vowel if glide insertion precedes the centralization of stressless non-high front vowels). This is shown by the forms in (111).¹¹⁰

84 *Phonological phenomena*

(111) <i>recEie</i>	[ɛ̃ji]	(<i>receEmos</i>	[jɛ̃])
‘(he) fier-SUBJ’		‘(we) fier-SUBJ’	
<i>pentEie</i>	[ɛ̃ji]	(<i>pentEmos</i>	[jɛ̃])
‘(he) comb-SUBJ’		‘(we) comb-SUBJ’	
<i>cEie</i>	[ɛ̃ji]	(<i>ceEmos</i>	[jɛ̃])
‘(he) sup-SUBJ’		‘(he) sup-SUBJ’	
<i>custEie</i>	[ɛ̃ji]	(<i>custEmos</i>	[jɛ̃])
‘(he) pay-SUBJ’		‘(he) pay-SUBJ’	

Finally, V2 has to be characterized as an oral vowel, in order to explain the possible absence of the glide in forms such as those in (112).¹¹¹

(112) <i>vEem</i>	[vɛ̃ʃ]	‘(they) see’
<i>dEem</i>	[dɛ̃ʃ]	‘(they) give’

Thus, the only restriction on V2 for which we have evidence is that it must be an oral vowel. In other words, the glide seems to be inserted between a stressed /e/ and any of the oral vowels that can appear after such a segment in EP. The following examples further illustrate this generalization.

(113) <i>custEio</i>	[ɛ̃ju]	‘(I) pay’	<i>mEio</i>	[ɛ̃ju]	‘half’
<i>custEia</i>	[ɛ̃jɐ]	‘(he) pays’	<i>mEia</i>	[ɛ̃jɐ]	‘half-FEM’
<i>custEie</i>	[ɛ̃ji]	‘(he) pay-SUBJ’			
(<i>custeAr</i>	[já]	‘(to) pay’)	(<i>meAdos</i>	[já]	‘middle’)

As shown in (114), glide insertion does not apply across words.

(114) <i>lê artigos</i>	[ɛ̃ɐ]/*[ɛ̃jɐ]	‘(he) reads articles’
<i>dê utensílios</i>	[ɛ̃u]/*[ɛ̃ju]	‘(he) give-SUBJ tools’

Since this process seems not to be constraint by morphological information and is not operative across words, we propose that glide insertion applies within the prosodic word domain, as in (115).

Glide insertion to break a hiatus appears to be a rather general process. It precedes another general (postlexical) process, which is V2 semivocalization (see section 11.1), since it bleeds it (i.e. glide insertion removes the context for V2 semivocalization, as for example /paseu/ becomes [pɛ̃sɛ̃ju]

instead of [pəsɛw]). In addition, it follows the morphologically conditioned phenomenon of word stress assignment. However, this gives us no information about its lexical/postlexical nature.

$$(115) \emptyset \rightarrow j / \left[\dots \begin{pmatrix} V_i \\ \text{-back} \\ \text{-high} \\ \text{-low} \end{pmatrix} - \begin{pmatrix} V \\ \text{-nasal} \end{pmatrix} \dots \right]_{\omega}$$

(where V_i is in a stressed position)

As we have seen by the forms in (109) above, when the segment that follows a stressed /e/ is a glide itself, glide insertion does not apply. However, the palatal glide surfaces between two underlying vowels even if the postlexical process of V1 semivocalization applies, causing the semivocalization of the first of two adjacent vowels (e.g. *passaio agradável* [pəsɛjwəgrədávɛɫ]/[pəsɛjwəgrədávɛɫ] ‘nice walk’ – see section 11.2 on V1 semivocalization). The fact that the [j] which results from glide insertion to break a hiatus still surfaces when the underlying vowel that follows /e/ is realized as [w] indicates that the process is no longer operative postlexically.

Additionally, glide insertion does not apply between a verb and an enclitic pronoun starting with a vowel, as it will be shown in chapter 4, section 3.2. This also suggests that the process at stake belongs to the lexical component. Indeed, the lexical nature of glide insertion straightforwardly explains why it does not apply when the context for its application results from the postlexical concatenation of words.

Finally, the glide is always indicated in the orthography, which also suggests the lexical nature of the rule.¹¹²

8. Lowering of stressless vowels in final syllables closed by /r/

Stressless vowels usually undergo vowel reduction. In some contexts, however, they systematically escape reduction, as we have pointed out in section 3. One such case is discussed in detail in this section.

As illustrated in (116), non-high stressless vowels in syllables closed by /r/ in word final position are obligatorily realized as low.

86 *Phonological phenomena*

- | | | | | | |
|--------------------|-----|-----------------|-----|---------------|-----|
| (116) <i>líder</i> | [ɛ] | <i>carÁcter</i> | [ɛ] | <i>jÚnior</i> | [ɔ] |
| ‘leader’ | | ‘character’ | | ‘junior’ | |
| <i>repÓrter</i> | [ɛ] | <i>revÓlver</i> | [ɛ] | <i>VÍtor</i> | [ɔ] |
| ‘journalist’ | | ‘pistol’ | | ‘id.’ | |

By contrast, high vowels surface as such, as shown in (117).

- | | | |
|--------------------|-----|----------|
| (117) <i>fÉmur</i> | [u] | ‘femur’ |
| <i>mÁrtir</i> | [i] | ‘martyr’ |

Since /ɛ/ and /e/, and /ɔ/ and /o/ are neutralized in this position and always surface as low, we propose that non-high vowels undergo a general process of lowering, which is responsible for this neutralization.

We should mention that stressless vowels in syllables closed by /l/ behave in a similar way, in that they also surface as low (see 118).¹¹³

- | | | | | | |
|----------------------|-----|-------------|---------------|-----|---------------|
| (118) <i>delgAdo</i> | [ɛ] | ‘thin’ | <i>palrAr</i> | [a] | ‘(to) babble’ |
| <i>relvInha</i> | [ɛ] | ‘grass-DIM’ | <i>Álcool</i> | [ɔ] | ‘alcohol’ |

The difference between the case of lowering under discussion and the one involving vowels in syllables closed by /l/, is that vowels in syllables closed by /r/ are only lowered in word final position, as shown in (119) – where the relevant vowels are in word internal position and undergo vowel reduction – whereas no such restriction applies to vowels in syllables closed by /l/ (see the examples in 118 above).¹¹⁴

- | | | |
|-----------------------|----------|-------------------|
| (119) <i>percebEr</i> | [i]/*[ɛ] | ‘(to) understand’ |
| <i>revertEr</i> | [i]/*[ɛ] | ‘(to) reverse’ |
| <i>mergulhAr</i> | [i]/*[ɛ] | ‘(to) dive’ |
| <i>cervEja</i> | [i]/*[ɛ] | ‘bier’ |

A second difference between this lowering process and the one applying to vowels in syllables closed by /l/ is its regular application also to round vowels (see 116 above), whereas round vowels in non-final syllables closed by /l/ are not necessarily low (see note 113 above).

This process of lowering seems to apply at the right-edge of the prosodic word domain. The fact that it applies to final syllables of stressed pre-

fixes and of non-final stems in root compounds, as shown in (120), suggests that the relevant domain is the prosodic rather than the morphological word (see also chapter 6, sections 3.2 and 3.3).

(120) <i>hIper-violento</i>	[ɛ]/*[i]	‘very aggressive’
<i>INter-urbano</i>	[ɛ]/*[i]	‘interurban’
<i>sUper-activo</i>	[ɛ]/*[i]	‘very dynamic’
<i>clber-café</i>	[ɛ]/*[i]	‘cyber coffee-shop’

The phenomenon under observation may actually affect any stressless vowel in a final syllable closed by a sonorant consonant. This seems true for syllables ending in [r], regardless of the quality of the relevant vowel (cf. *revólver* ‘pistol’, *júnior* ‘junior’). It is also true for syllables closed by /l/ (cf. *nível* ‘level’), even with round vowels (cf. *álcool* ‘alcohol’). And the same also happens with vowels in final syllables exceptionally closed by a nasal consonant (cf. *sémen* ‘semen’, *cólofon* ‘colophon’). The latter cases are particularly interesting in this respect, because EP has a rather general process that causes low vowels to surface as mid before nasal consonants. Thus, the presence of a low vowel in these contexts strongly suggests that the process of lowering also affects final syllables closed by a nasal consonant.

We therefore propose to account for this lowering process by (121).¹¹⁵

$$(121) [-\text{high}] \rightarrow [+low] / [\dots - \left(\begin{array}{c} C \\ +\text{sonorant} \end{array} \right) (C)]_{\omega}$$

(where the vowel belongs to a stressless syllable)

Here we should make a brief digression in order to motivate the inclusion of an optional final consonant in our rule. The optional consonant is meant to allow for the forms under discussion to end in fricative, since when words ending in [r] appear in the plural form, the vowel is still lowered, as in (122).

(122) <i>líderes</i>	[ɛrʃ]	‘leaders’
<i>repórteres</i>	[ɛrʃ]	‘journalists’

The underlying form of the plural suffix has been a matter of some controversy in the literature. We are adopting here the analyses of d'Andrade (1977)¹¹⁶ and Morales-Front and Holt (1997), where it is proposed that the underlying form of the plural suffix in EP is /s/. In fact, if it is assumed instead that there is an underlying vowel, that is, that the form of the plural suffix is /es/ (cf. Mateus 1975), then the plural forms shown in the first column of (123) violate the three syllable stress window, otherwise totally pervasive in EP: although the norm imposes the stress to shift in these forms (against the generalization that non-verbal inflectional suffixes do not affect stress placement in EP), for most speakers the stress is kept in its original position, thus allowing for the stress to appear in what would count as the fourth syllable from the right.

(123) Generalized form	Norm	
<i>jUniores</i>	<i>juniOres</i>	'junior-PL'
<i>sEniores</i>	<i>seniOres</i>	'senior-PL'

Thus, if pluralization added an extra syllable, the cases with no stress shift would contradict the generalization that the three syllable stress window is universal (e.g. Van der Hulst 1996). We take these facts as evidence that the sequence tap plus fricative is allowed in word final syllables in EP.

In addition to this, words such as those in (124) show that this sequence is also allowed in word internal position.

(124) <i>perspectIva</i>	'perspective'
<i>perspicAz</i>	'acute'

Contrary to what is predicted by Mateus and d'Andrade's (2000: chap.3) proposal on syllable construction, we believe that this consonant cluster may be included in the same syllable.¹¹⁷ We thus assume that this cluster is possible in word final position, in the same way it is allowed in other positions of the word.¹¹⁸

To conclude this brief detour, we should notice that for some speakers, a schwa may surface instead of a front low vowel, both in the singular and in the plural forms of the word *líder* 'leader'. However, in this case, the relevant syllable is in fact not in word final position because a schwa usually surfaces after the tap (i.e. *lid(e)re(s)*).¹¹⁹ What these facts seem to suggest is that, for some speakers, the underlying form of this word contains an underlying final vowel (like in *livre* 'free'). Consequently, here

lowering does not apply because the relevant vowel is simply not in a closed syllable in word final position. The fact that this does not happen with other words ending in [r], like *repórter* ‘journalist’, also supports our view.

As we have said above, this process applies regularly. Its productivity may be seen in the adaptation of loan words, as in (125).

- (125) *pulÔver* [ɛ] ‘pullover’ *poliÊster* [ɛ] ‘a kind of tissue’
 sÊter [ɛ] ‘setter’ *ROver* [ɛ] ‘id’

Nevertheless, the fact that the vowel is low even if the tap is assigned to a following empty onset, as a consequence of phrasal resyllabification, supports the lexical application of the process under discussion. This is illustrated in (126).

- (126) *o repÔrter acha* (o.re.por.te.ra.cha) [ɛ]/*[i]
 ‘the journalist thinks’

9. Initial /r/-strengthening

There are two possible realizations for /r/ that depend on the context: an alveolar tap ([ɾ]) is found in intervocalic position, in the second position of complex onsets and in coda (see 127a); a voiced uvular trill ([R]) or fricative ([ʁ]) is found in syllable initial position if preceded by a consonant in the coda of the preceding syllable, and in word initial position (see 127b).¹²⁰

- (127) a. *caro* [ɾ] ‘expensive’ b. *melro* [R] ‘blackbird’
 prato [ɾ] ‘dish’ *tenro* [R] ‘tender’
 carga [ɾ] ‘load’ *Israel* [R] ‘id.’
 rato [R] ‘mouse’

Mateus and d’Andrade (2000: 2.4.2) propose that /r/ is underlyingly underspecified for the place features, and that a default rule specifies it as coronal, while an “onset rule” specifies it as [+back], thus yielding the realizations [ɾ] and [R], respectively.¹²¹

The generalization that the uvular trill occurs either in word initial position or in a syllable onset preceded by a filled coda, suggests that some process has indeed applied changing (or specifying) /r/ into [R]. This analysis explains why [r] is the only consonant allowed in coda that cannot precede [R], whereas the other consonants can, as shown in (127b) above. It explains furthermore why it does not occur in coda only when followed by another [r], whereas it can appear with most EP consonants (see 128).¹²²

(128)	<i>ha[rp]a</i>	<i>a[rb]orizar</i>	<i>a[rm]a</i>	<i>pé[rf]ido</i>	<i>a[rv]oredo</i>
	'harp'	'(to) forest'	'weapon'	'nasty'	'grove of trees'
	<i>a[rt]e</i>	<i>a[rd]or</i>	<i>a[rs]ebispo</i>	<i>A[rz]ebeijão</i>	<i>ca[rn]e</i>
	'art'	'ardour'	'archbishop'	'Azerbaijan'	'meat'
	<i>o[rl]a</i>	<i>a[rʃ]ote</i>	<i>A[rʒ]entina</i>	<i>a[rk]o</i>	<i>espa[rg]o</i>
	'border'	'torch'	'id'	'arc'	'asparagus'

As pointed out by Mateus and d'Andrade (2000: 2.2.1.2), this analysis also captures the absence of the uvular trill at the onset of final syllables in words with antepenultimate stress. Since in EP there are no closed/heavy syllables in post-stressed non-final positions, no syllable may have an /r/ in coda in such position, and thus no context is provided for /r/-strengthening in the final syllables of proparoxiton words (see 129).¹²³

(129)	<i>fósforo</i>	[r]	<i>*fósforro</i>	[R]	(*fós.for.ro)	'match'
	<i>pífaro</i>	[r]	<i>*pífarro</i>	[R]	(*pí.far.ro)	'a flutelike instrument'
	<i>pólvora</i>	[r]	<i>*pólvorra</i>	[R]	(*pól.vor.ra)	'powder'

In Mateus and d'Andrade's account of the process, no explicit distinction is made between the context of the rule that is responsible for the uvular realization in syllable initial position preceded by a consonant and the uvular realization that occurs in word initial position. Nevertheless, it seems that the two cases should be set apart, for there is no obvious relation between the two contexts of /r/-strengthening. We therefore propose that in the latter case the uvular realization is a strengthening process that is dependent on word initial position, as expressed in (130).¹²⁴

$$(130) \text{r} \rightarrow \text{R} / [_ \dots]_{\omega}$$

This process has no exceptions. In addition, its productivity can be seen in Portuguese adaptations of English words, even by speakers that have English as a second language.¹²⁵ This is illustrated in (131).¹²⁶

- (131) *all* [R]ight (children's adaptation)
 [R]obe[r]t [R]edfo[r]d (general)
 Ma[r]ilyn Mon[R]oe (general)
 [R]ent a ca[r] (general)
 p[r]ime [R]ate (advertising spot)

Despite its wide application, /r/-strengthening seems to operate in the lexical component, as the following facts suggest.

It precedes the deletion of the syllable final nasal segment, as shown by forms such as *tenro* [tẽru]: the nasal segment is linked to the nucleus of the syllable, and thus no segment fills the coda that precedes /r/.¹²⁷ Notice that the composition of the nucleus preceding a /r/ in syllable initial position does not affect its realization. For instance, a branching nucleus does not trigger the process, as shown in (132).¹²⁸

- (132) *Cairo* [r] 'id.' *eira* [r] 'threshing floor'
 Laura [r] 'id.' *loiro* [r] 'blond'

Since nasal resolution must be accomplished within the lexical component (see section 5), this process must also be lexical. Moreover, its non-application to a tap that becomes prosodic word initial due to postlexical resyllabification, as in (133), also argues for its lexical status.

- (133) *cantar alegre* (can.ta.ra.le.gre) [r]/*[R] 'singing joyful'

As we will see in chapter 5, section 3, the adjunction of prefixes with open syllables to a base starting with /r/ does not prevent this segment from undergoing /r/-strengthening. This will be taken to show that prefixes are not incorporated into the prosodic word that dominates the morphological base.

10. The realization of word initial vowels

Word initial non-high vowels are realized with different degrees of opening in a fashion that seems almost unruly: a front vowel that surfaces in stressed position as low is realized as mid or high in stressless initial position in some cases (see 134a), while in other cases it is obligatorily realized as low (see 134a'); in some cases the same underlying central vowel is realized as low (see 134b) or as mid (see 134b'); in some forms /ɔ/ is realized obligatorily as low (see 134c) and in others it may be realized as mid or low (see 134c').

(134) a. ([ɛ]dito 'edict')	<i>editAl</i>	*[ɛ]/[e]/[i]	'decree'
a'. ([ɛ]mico 'ethnic')	<i>etnologia</i>	[ɛ]/*[e]/*[i]	'ethnology'
b.	<i>activIsta</i>	[a]/*[ɐ]	'activist'
b'.	<i>actividAde</i>	*[a]/[ɐ]	'activity'
c. ([ɔ]ptimo 'great')	<i>optimIsta</i>	[ɔ]/*[o]	'optimistic'
c'. ([ɔ]ra '(he) prays')	<i>orAr</i>	[ɔ]/[o]	'(to) pray'

In addition to this, there is great dialectal and interspeaker variation.¹²⁹ To illustrate these differences, (i) in Northern dialects an initial low round vowel may surface as high – for example, in *orar* 'to pray' the initial vowel may be realized as [u] – whereas in other dialects the vowel must be low or mid; (ii) an initial vowel that is obligatorily low for some speakers, may surface as mid for other speakers – for example, contrary to the realizations given in (134), *activista* for some speakers may present an initial [ɐ], in the same way as *atividade* may be realized by other speakers with an initial [a].

A further source of difficulty in the analysis of this phenomenon is that while the variation in height suggests that vowels are specified for the height features only postlexically, the realization of initial vowels seems to be also conditioned by lexical information.

All these facts taken together render the realization of initial vowels a difficult phenomenon to capture. In fact, its complexity suggests that it is possibly the result of a change in progress, which has not yet reached its stability.¹³⁰

Our approach in this section, as throughout this book, is to concentrate only on data belonging to a single dialect. In particular, we use introspec-

tion data in order to try to capture the regularities that may be found within a single system.

As we have seen in section 3, vowels do not undergo (full) vowel reduction in word initial position. Specifically, initial front vowels never surface as [i] and low round vowels are not raised to [u]. Nevertheless, stressless vowels in this position are also subject to a raising process. Contrary to vowels in non-initial position, in this case each type of vowel seems to rise just one degree.¹³¹ In the following examples, we present one of the possible realizations of each type of initial vowel (other realizations are discussed below).¹³²

(135) /e/ ([é]rro)	<i>errAr</i>	[i]/*[i]	([é]rgo)	<i>erguer</i>	[i]/*[i]
‘mistake’	‘(to) mistake’		‘(I) raise’	‘(to) raise’	
/ε/ ([é]rdo)	<i>herdAr</i>	[e]/*[i]	([é]co)	<i>ecoar</i>	[e]/*[i]
‘(I) inherit’	‘(to) inherit’		‘echo’	‘to echo’	
/a/ ([á]vido)	<i>avidEz</i>	[ɐ]	([á]rte)	<i>artIsta</i>	[ɐ]
‘avid’	‘avidity’		‘art’	‘artist’	
/o/ ([ó]vo)	<i>ovIto</i>	[u]	([ó]lho)	<i>olhIto</i>	[u]
‘egg’	‘egg-DIM’		‘eye’	‘eye-DIM’	
/ɔ/ ([ó]ra)	<i>orAr</i>	[o]/*[u]	(h[ó]ra)	<i>horÁrio</i>	[o]/*[u]
‘(he) prays’	‘(to) pray’		‘time’	‘timetable’	

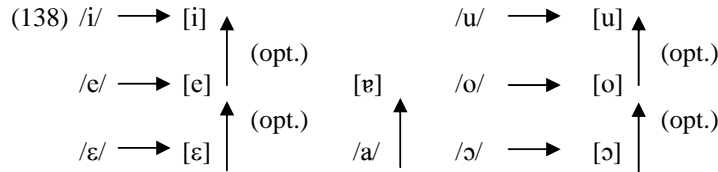
Non-central low vowels may surface either as mid or low, and mid vowels may surface either as high or mid (see 136).

(136) /ε/	([é]rdo)	<i>herdAr</i>	[ε]/[e]	([é]co)	<i>ecoar</i>	[ε]/[e]
	‘(I) inherit’	‘(to) inherit’		‘echo’	‘to echo’	
/ɔ/	([ó]ra)	<i>orAr</i>	[ɔ]/[o]	(h[ó]ra)	<i>horÁrio</i>	[ɔ]/[o]
	‘(he) prays’	‘(to) pray’		‘time’	‘timetable’	
/e/	([é]rro)	<i>errAr</i>	[e]/[i]	([é]rgo)	<i>erguer</i>	[e]/[i]
	‘mistake’	‘(to) fail’		‘(I) raise’	‘(to) raise’	
/o/	([ó]vo)	<i>ovIto</i>	[u]/[o]	([ó]lho)	<i>olhIto</i>	[u]/[o]
	‘egg’	‘egg-DIM’		‘eye’	‘eye-DIM’	

Initial high vowels do not show any variation, as illustrated in (137).

(137) <i>iIStre</i>	[i]	‘illustrious’	<i>usado</i>	[u]	‘used’
<i>ilhÉu</i>	[i]	‘islander’	<i>utilizAr</i>	[u]	‘(to) employ’

The figure in (138) summarizes these facts.



As for the raising in one degree of the central vowel, we propose to account for it by the rule in (139), which applies obligatorily.¹³³

$$(139) \begin{pmatrix} \text{-round} \\ \text{+back} \end{pmatrix} \rightarrow [-\text{low}] / [_ \dots]_{\omega}$$

(where the vowel is stressless)

This rule, like the more general rule of vowel reduction, has exceptions, as those in (140).

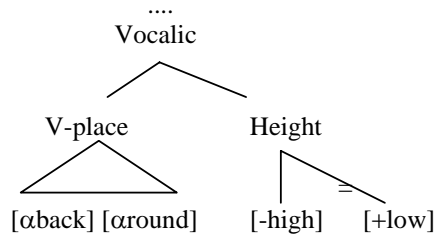
- (140) *activIsta* [a]/*[ø] 'activist'
accionista [a]/*[ø] 'stockholder'

As for the other cases where two alternative realizations are possible, there are two basic facts that we would like to capture: the neutralization in height features and the optionality in the realization of those features. The analysis we develop in the following paragraphs is an attempt in this direction.

With respect to non-high non-central vowels that are not underlyingly specified as mid, we propose that they are underspecified for the low feature. In addition, underspecification may be either underlying, or it may result from the application of a rule that eliminates the [low] feature. We will use feature geometry and the autosegmental framework in order to analyze this process as the delinking of the [low] feature, as in (141).¹³⁴ We propose that this delinking process is obligatory and operates in the lexicon, and that the unspecified low feature is only specified postlexically. Thus, the initial vowel of words like *herdar* or *horário* becomes lexically underspecified for the [low] feature and this feature is specified postlexically as

plus or minus, so that it may vary in height. Therefore, both the realizations [ɛ]/[e] and [o]/[ɔ], respectively, are possible.

(141) Delinking of [low] in word initial position (non-central low vowels)



(where the vowel is stressless and in prosodic word initial position)

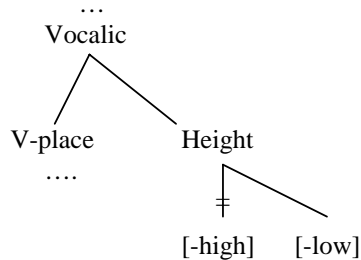
According to this analysis, the optionality in the realization of these vowels as mid or low follows from a postlexical process of feature specification.¹³⁵ This approach allows, furthermore, for some forms to be lexically marked in order not to undergo delinking, like those in (142). Hence, for these forms there is no variation in the realization of the relevant vowels.

(142) *etnologia* [ɛ]/*[e] 'ethnology'
optimista [ɔ]/*[o] 'optimistic'

The same type of analysis can also be extended to mid vowels. The operation of delinking applies to vowels specified as [-low], but in this case delinking affects the feature [high]. The process is represented in (143). Thus, the initial vowel of words like *errado* or *olhito* becomes lexically underspecified for the feature [high] and this feature is postlexically specified as plus or minus, so that its realization alternates between [i] and [e], or [u] and [o], respectively.

Under this approach, we may explain, for instance, why the forms in (144a) may show the realizations [e] or [i], while the forms in (144b) may show the realizations [ɛ] or [e]. We simply have to admit that in the former cases the underlying vowel is /e/ (or, at least, underspecified for the low feature) while in the latter it is /ɛ/.

(143) Delinking of [high] in word initial position (mid vowels)



(where the vowel is stressless and in prosodic word initial position)

- (144) a. *ervIlha* [e]/[i]/*[ε] 'pea' b. *ErnEsto* [ε]/[e]/*[i] 'id.'
- ermIda* [e]/[i]/*[ε] 'hermitage' *HerculAno* [ε]/[e]/*[i] 'id.'

Comment [SFeMV1]:

To summarize, the neutralization in height of initial vowels is proposed to be due to a generalized process of delinking of height features that operates in the lexical component. The variation in the realization of initial vowels, by contrast, may be seen to be the result of the variability in the postlexical specification of the unspecified features.¹³⁶

To reinforce the analysis just presented, we should stress that, besides words starting with a central vowel (which is necessarily an underlying low vowel), for most words starting with non-high vowels speakers have no cues as to the height of the underlying segment. In those cases where the alternation in the realization of vowels is possible, we find in general the alternation between [i]/[e] and [o]/[ɔ] (see 145). This suggests the lack of lexical definition of the values for the features [high] and [low] of non-back and of round vowels, respectively.

- | | | | |
|----------------------|-----------------|-------------------|--------------|
| (145) <i>equador</i> | <i>erosão</i> | <i>equilíbrio</i> | [i]/[e]/*[ε] |
| 'equator' | 'erosion' | 'balance' | |
| <i>educar</i> | <i>elefante</i> | <i>efeito</i> | [i]/[e]/*[ε] |
| '(to) teach' | 'elephant' | 'effect' | |
| <i>evasão</i> | <i>evidente</i> | <i>economia</i> | [i]/[e]/*[ε] |
| 'escape' | 'evident' | 'economy' | |

<i>ocupar</i>	<i>odor</i>	<i>opinião</i>	*[u]/[o]/[ɔ]
‘(to) occupy’	‘odor’	‘opinion’	
<i>objetivo</i>	<i>omisso</i>	<i>opressão</i>	*[u]/[o]/[ɔ]
‘goal’	‘silent’	‘oppression’	
<i>obedecer</i>	<i>ofuscar</i>	<i>oportunidade</i>	*[u]/[o]/[ɔ]
‘(to) obey’	‘(to) obscure’	‘opportunity’	

As for the exceptional vowels that may only surface either as low or as mid, they must be considered instances of the exceptional non-application of the processes involved in initial vowels realization. These segments must be marked in order not to undergo (obligatory) central vowel raising and/or the operations of delinking of the features [low]/[high].

Both central vowel raising and delinking appear to be special cases of vowel reduction. In fact, these processes fail to apply (that is, there is no variation in the realization of initial vowels) in the same circumstances as vowel reduction does not apply: variation is not possible in initial vowels belonging to syllables closed by /l/ or containing diphthongs, as well as in unpredictable exceptional cases, as we have seen in the previous paragraph (see also 146).

(146) <i>helvético</i>	[ɛ]/*[e]	<i>almoçar</i>	[a]/*[ɐ]		
‘Helvetian’		‘(to) lunch’			
<i>Europa</i>	[e]/*[i]	<i>aumentar</i>	[a]/*[ɐ]	<i>oitavo</i>	[o]/*[u]
‘Europe’		‘(to) enlarge’		‘eighth’	
<i>etnologia</i>	[ɛ]/*[e]	<i>ativista</i>	[a]/*[ɐ]	<i>optimista</i>	[ɔ]/*[u]
‘ethnology’		‘activist’		‘optimistic’	

Formerly, we have attempted to explain the low realization of round vowels as following from a general rule of lowering of mid vowels (Vigário 2000a). The existence of mid round vowels that cannot surface as low, however, made us abandon this analysis (cf. *ovito* ‘egg-DIM’, *oquinho* ‘empty-DIM’). Nevertheless, the fact that there is optionality in the realization of initial vowels has led us to propose in this section an analysis that may account for this as a postlexical phenomenon. Given that optionality is usually a symptom of a postlexical operation, if feature specification is accomplished postlexically, the aspects related to the variation in height described here are indeed accounted for by a postlexical operation.

Although we do not have direct evidence for the morphological or prosodic nature of the initial position involved in this process, the fact that

initial vowels are never as reduced as non-initial vowels may be thought to be a consequence of initial position strengthening, which may be viewed as a prosodic property.

An alternative analysis is found in Mateus (1995) and Mateus and d'Andrade (2000: 3.3.1), who have proposed to account for the absence of centralization of stressless non-back vowels and of raising of underlying /o/ and /ɔ/ by resorting to syllable structure. According to Mateus and d'Andrade, it is the empty onset of word initial syllables that prevents the occurrence of initial [i] and [u], when they result from the application of vowel reduction.

We depart from this view for several reasons. First, we have seen that an underlying /o/ may be raised to [u] in initial onsetless syllables (like in *olhito* in 135 above). Thus, in this case, the empty onset did not block vowel raising. It is not clear, furthermore, why raising to [i] is allowed, while raising to [u] is not. In addition, while the segmental make-up of syllables may affect the quality of stressless vowels, in EP this only happens with segments that are part of the rhyme: vowels in branching nuclei are not reduced, neither are nasalized vowels, nor vowels in syllables closed by /l/ or in final syllables closed by sonorant consonants. In other words, the onset composition never affects vowel reduction. The final and most compelling argument against the view that vowels in onsetless syllables do not reduce comes from the behavior of vowels in similar conditions in word internal position. As illustrated in (147), vowels in onsetless syllables do reduce in word non-initial position.

- (147) (*mi*[ó]*lo* 'soft part of bread') *mioll**inho* [u] 'soft part of bread-DIM'
 (*ca*[ɔ]*tico* 'chaotic') *caotiz**Ar* [u] '(to) make chaotic'
 (*ado*[ɛ]*ce* '(he) becomes sick') *adoec**Er* [i] '(to) become sick'

By contrast, the hypothesis that vowel reduction is not as extreme word initially as it is within the word because this position is prosodically strong (see also Andrade 1980) is consistent with other phonological phenomena that characterize this position and that can be related to its strength: namely, /r/-strengthening, initial stress and emphatic stress assignment (see, respectively, sections 9, 16, and 17). There is, in addition, at least another case where stressless vowels in a specific position of the prosodic word are not subject to vowel reduction – when vowels occur in syllables closed by a sonorant consonant in *prosodic word final position* (see section 8). Finally,

while the relation between a strong position and (less) vowel reduction is a motivated one, since reduction is characteristic of elements in weak positions, the relation between empty onsets and vowel reduction is not entirely transparent to us.

Putting aside the question of what the best analysis for the facts reviewed in this section is, we will assume that: (i) the variation in initial vowel height is only possible in prosodic word initial position; (ii) if a vowel reduces more than it is allowed in initial position, this means that the vowel cannot be considered to be in prosodic word initial position; (iii) the absence of total vowel reduction in initial position is related to the *strength* of the left-edge of the prosodic word domain in EP.

As we will see in chapter 5, the same type of variation in initial vowels is also found when words are subject to prefixation. Thus, vowels preceded by prefixes behave like initial vowels, even if resyllabification causes the onset of their syllables to be filled. This will be taken to indicate that prefixes are not phonologically treated as units incorporated into the prosodic word. And the same behavior is also found with prosodic words preceded by proclitics.

11. Semivocalization

There are two possible types of diphthongs in EP: falling diphthongs, of the form VG, and rising diphthongs, of the form GV. D'Andrade (1977: 28) formalizes in a single rule, presented in (148), the semivocalization of the first of two adjacent vowels (V1) and the semivocalization of the second of two adjacent vowels (V2).

$$(148) \begin{pmatrix} +\text{high} \\ -\text{stressed} \end{pmatrix} \rightarrow [-\text{syllabic}] / \text{ V}$$

However, as we will see in the next paragraphs, the environment for the two types of semivocalization is different, thus suggesting that two distinct processes are operative in EP.

11.1. V2 semivocalization (>VG)

Although Mateus and d'Andrade (2000) propose that surface glides are always underlying vowels, we will try to show that there are two possible sources for falling diphthongs in EP. In fact, in our view a distinction needs to be established between those cases where a glide *never* alternates with a vowel (see 149), and those cases where there is some optionality in the realization of a high vowel as a glide (see 150) (see also d'Andrade and Viana 1994).¹³⁷

(149) <i>viu</i>	[íw]	<i>pneu</i>	[éw]	<i>pai</i>	[áj]
'(he) saw'		'tyre'		'father'	
<i>chapéu</i>	[éw]	<i>mau</i>	[áw]	<i>sai</i>	[áj]
'hat'		'bad'		'(he) leaves'	

- (150) a. *rio* [íu]/[íw] 'river'
tio [íu]/[íw] 'uncle'
 (–o is a class-marker also found in *belo* 'beautiful' or *carro* 'car')
- b. *rio* [íu]/[íw] '(I) laugh'
doo [óu]/[ów] '(I) donate'
 (–o is a person-number suffix also found in *falo* 'I speak' or *canto* 'I sing')

Thus, regardless of the exact way the glide is obtained in the forms in (149), there is evidence that it is different from the semivocalization process illustrated in (150).¹³⁸ In addition, V2 semivocalization seems to operate automatically only in word final position, as shown by the absence of a glide in the words in (151) (the forms inside parentheses in the right column show the relevant vowel in stressed position in morphologically related words, thus demonstrating its underlying vocalic status).

(151) <i>saidlnha</i>	[ɐi]/*[áj]	'exit-DIM'	(<i>saÍda</i> 'exit')
<i>paułlsmo</i>	[ɐu]/*[aw]	'a literary style'	(<i>paUl</i> 'swamp')
<i>graudlnho</i>	[ɐu]/*[aw]	'big-DIM'	(<i>graÚdo</i> 'big')
<i>miudlnho</i>	[íu]/*[íw]	'small-DIM'	(<i>miÚdo</i> 'small')

Note that vowel reduction applies to the central vowel, a fact that further indicates that there is no diphthong, as vowel reduction regularly does not affect diphthongs (see section 3 above).¹³⁹

A further difference between V1 semivocalization and V2 semivocalization, is that, contrary to V1 semivocalization (see section 11.2 below), only round vowels appear to regularly undergo V2 semivocalization, as shown by the data in (152).¹⁴⁰

- | | | | |
|---------------------------|--------|-------------------|--------|
| (152) <i>amUe</i> | ??[új] | <i>vOe</i> | ??[ój] |
| ‘(he) become sullen-SUBJ’ | | ‘(he) fly-SUBJ’ | |
| <i>actUe</i> | ??[új] | <i>sOe</i> | ??[ój] |
| ‘(he) act-SUBJ’ | | ‘(it) sound-SUBJ’ | |

Also unlike V1 semivocalization, there is no V2 semivocalization if the two vowels do not belong to the same word, as shown in (153).

- | | | | | | |
|-------------------------|-------|-------------------------|-------|---------------------|-------|
| (153) <i>tÁxi usAdo</i> | *[iw] | <i>dÁ idEia</i> | *[áj] | <i>cAso isolAdo</i> | *[uj] |
| ‘old taxi’ | | ‘(it) gives (the) idea’ | | ‘isolated case’ | |
| <i>café imitAdo</i> | *[éj] | <i>cAsa humilde</i> | *[ɐw] | <i>avÓ idOsa</i> | *[ɔj] |
| ‘copied coffee’ | | ‘humble house’ | | ‘old grandmother’ | |

We should notice that, although these facts correspond to those already reported in Gonçalves Viana (1883), the description provided in Sá Nogueira (1938) shows some differences. Considering the realization of adjacent vowels across words, the first of which (V1) is stressed and the second (V2) is high and stressless, V2 is reported to be realized either as a vowel or as glide. However, the possibility of semivocalization is seen as dependent on the vowel quality of V1, since V2 semivocalization is possible if V1 is *i* or *a*, but not *e* or *o*. This is not consistent with our own intuitions, for, as far as we can judge, in our own dialect only a vowel may surface in this environment.

Notice that Sá Nogueira does not seem coherent, since between stressless vowels in the same conditions, no semivocalization of V2 is reported to be possible. In the latter case, Gonçalves Viana (1883) provides the same sort of data, with which we also agree. Finally, V1 high *stressed* vowels are also described in Sá Nogueira to optionally undergo semivocalization. Again, our dialect differs. In fact, the semivocalization of stressed vowels in similar environments was experimentally shown not to occur in present-day speech of the Lisbon variety of EP in Frota (2000).¹⁴¹ We

102 *Phonological phenomena*

therefore conclude that Sá Nogueira's observations do not apply to the system we are describing here, and that, in our phonological system, regular V2 semivocalization is only possible in word final position.

The fact that this process may also affect enclitic pronouns (e.g. *li-o* 'I read it') shows that the relevant notion of word is not morphological, but rather phonological (see chapter 5, section 4.3.1). Thus, we can establish that the process under discussion is the result of an optional prosodic word limit rule, as represented in (154).

$$(154) u \rightarrow w / [\dots V _]_{\omega}$$

To conclude this section, it should be noticed that V2 semivocalization has the properties of a postlexical phenomenon: (i) it is general; (ii) it applies optionally; (iii) and it may affect enclitics, which we will argue in chapters 4 and 5 to be postlexically incorporated into the prosodic word that includes their host.

11.2. *V1 semivocalization (>GV)*

The process that creates rising diphthongs affects stressless high vowels when followed by another vowel. This process is characterized in Frota (2000) as a span rule bounded by the intonational phrase domain. That is, it applies whenever its segmental context is met, within the intonational phrase domain. This is shown in (155), from Frota (2000: 84).¹⁴²

$$(155) [+high] \rightarrow [-syllabic] / [\dots _ V \dots]_{I}^{\max}$$

The application of this rule is illustrated in (156).¹⁴³

(156) a.	<i>iAto</i> 'hiatus'	[j]	<i>pass<u>e</u>Ar</i> '(to) walk'	[j]	<i>act<u>u</u>Ar</i> '(to) act'	[w]
	<i>família</i> 'family'	[j]	<i>are<u>A</u>l</i> 'sand field'	[j]	<i>contínu<u>a</u></i> 'continuous-FEM'	[w]
b.	<i>p<u>E</u>de-o</i> '(he) asks-for it'	[j]	<i>qu<u>E</u>ro-a</i> '(I) want-it-FEM'	[w]		
	<i>g<u>O</u>sto de anim<u>A</u>is</i> '(I'm) fond of animals'	[j]	<i>g<u>O</u>sto do an<u>E</u>l</i> '(I'm) fond of-the ring'	[w]		

c. <i>grANde Área</i>	[j]	<i>sAlto Alto</i>	[w]
‘penalty area’		‘high heel’	

Thus, in the case of V1 semivocalization, not only high round vowels but also high non-round vowels may semivocalize.

Our characterization of V2 semivocalization as a domain limit rule, as well as the span nature of V1 semivocalization are consistent with the realization of a sequence of two adjacent unstressed high vowels ([iu] or [ui]) within the word. As shown in (157), in this environment only V1 may undergo semivocalization.¹⁴⁴

(157) <i>ruído</i>	[wi]/*[uj]	
‘noisy’	([i] is an underlying vowel, as in <i>ruído</i> ‘noise’)	
<i>suinicultura</i>	[wi]/*[uj]	
‘pig breeding’	([i] is an underlying vowel, as in <i>suíno</i> ‘pig’)	
<i>miudezas</i>	[ju]/*[iw]	
‘little things’	([u] is an underlying vowel, as in <i>miúdo</i> ‘small’)	

The only case where the context for both rules is met is in prosodic word final position. The forms in (158) show that, in these cases, only V1 semivocalization is possible.

(158) <i>armário</i>	[ju]/*[iw]	‘closet’
<i>homogéneo</i>	[ju]/*[iw]	‘homogeneous’
([u] is an underlying vowel – the class marker –o)		

Thus, the application of V1 semivocalization wins over V2 semivocalization, for reasons we leave unexplored here.

Like V2 semivocalization, V1 semivocalization presents the properties that allow us to classify it as a postlexical process: (i) it does not refer to morphological or lexical information; (ii) its context of application may be obtained as a consequence of the (postlexical) concatenation of words; (iii) and it is an optional process.

12. Final non-back vowel deletion

Non-back stressless vowels in word final position may surface either as a [i] or as a schwa. [i] is found in words ending in an underlying /i/ that do not belong to the verbal class, as shown in (159).¹⁴⁵

- (159) *tÁxi* [i] 'taxi' *jÚri* [i] 'jury'
 dÉrbi [i] 'special game' *dÂndi* [i] 'dandy'

As for non-high non-back vowels, they regularly surface as a high central vowel ([ɨ]), which is in fact usually deleted (see 160).

- (160) *tenEnte* ([ɨ])/0 'tenant' *bAlde* ([ɨ])/0 'pail'
 bEbe ([ɨ])/0 '(he) drinks' *fAle* ([ɨ])/0 '(he) speak-SUBJ'

For the sake of exposition we have referred to these final vowels as non-back vowels in previous work (Vigário 1998b, 1999a). However, we do not have clear-cut evidence on whether the postlexical processes that affect these vowels (namely, vowel deletion and semivocalization, to be discussed below) operate before or after the specification of underlying non-high non-back vowels as central. In this section, we refer to these final segments either as *non-back* or *schwa* vowels.¹⁴⁶

As we have said, although schwa may surface as such in absolute final position, it is usually deleted. The same is true if it is followed by a word starting with consonant, as shown in (161).

- (161) a. *forte* 0 b. *forte pessoa* 0
 'strong' 'strong person'
 pobre 0 *pobre mulher* 0
 'poor' 'poor woman'
 deve 0 *deve saber* 0
 '(he) should' '(he) should know'

If a word ending in schwa is followed by a word starting with a vowel, the schwa may either be deleted or, under certain circumstances, surface as a palatal glide. Partial descriptions of the realization of schwa in this environment can be found in Gonçalves Viana (1883), Sá Nogueira (1939), d'Andrade and Viana (1993), Ellison and Viana (1996), and Vigário (1998b). As already noted in the latter work, these descriptions show a

number of differences, which we believe to reflect a phonological change.¹⁴⁷ In what follows, we will focus on the facts of our own system, which we assume to represent the present stage of the process in the EP variety of speakers belonging to the same sociolinguistic group as ourselves.¹⁴⁸

Starting with the most common situation, schwa is usually deleted if followed by a vowel (V2), whether stressed or unstressed, as shown in (162) (here and elsewhere in this book, we use the symbol ‘^x’ to indicate that the signaled realization is nearly impossible).

- | | | | |
|------------------------------|---------------------|---------------------|---------------------|
| (162) a. <i>forte abrIgo</i> | 0/ ^x [j] | b. <i>pele Alva</i> | 0/ ^x [j] |
| ‘strong shelter’ | | ‘white skin’ | |
| <i>pobre artIsta</i> | 0/ ^x [j] | <i>enorme Obra</i> | 0/ ^x [j] |
| ‘poor artist’ | | ‘huge work’ | |
| <i>deve optAr</i> | 0/ ^x [j] | <i>verde Único</i> | 0/ ^x [j] |
| ‘(he) should choose’ | | ‘green unique’ | |

Stress in V2 plays, however, an important role in blocking schwa deletion, but only in compound-like units, as in (163) (see chapter 6 for the analysis of this type of structures).

- | | | |
|---------------------------|--------|---------------------------|
| (163) <i>grANde Área</i> | *0/[j] | ‘penalty area’ |
| <i>PSR (pÊ Esse Erre)</i> | *0/[j] | ‘abbreviation of the name |
| | | of a political party’ |
| <i>dIr-te-Emos</i> | *0/[j] | ‘(we) shall tell-you’ |

Some forms with numerals also pattern as compound structures, as in (164a), while others do not (164b).

- | | | | |
|----------------------------|--------|--------------------------|---------------------|
| (164) a. <i>onze hOras</i> | *0/[j] | b. <i>dezanove hOras</i> | 0/ ^x [j] |
| ‘eleven hours’ | | ‘nineteen hours’ | |
| <i>doze hOras</i> | *0/[j] | <i>sete hOras</i> | 0/ ^x [j] |
| ‘twelve hours’ | | ‘seven hours’ | |
| <i>doze Anos</i> | *0/[j] | <i>dezassete Anos</i> | 0/ ^x [j] |
| ‘twelve years’ | | ‘seventeen years’ | |

As it will be shown in chapter 6, section 3.7, the reference to the specific numeral is not enough to account for the absence of the glide, since the same numerals followed by different words starting with a stressed vowel

106 *Phonological phenomena*

show the same behavior as other non-compound word sequences. This can be seen by the contrast between the sequences in (164a) and (165).

- (165) *onze Óleos* 0/^x[j] ‘eleven oils’
 doze Ovos 0/^x[j] ‘twelve eggs’
 doze Ânforas 0/^x[j] ‘twelve amphoras’

Other cases where deletion is impossible involve pronominal clitics. However, once again the reference to the category of pronouns is not enough to account for the presence/absence of the glide, as shown in (166). Both in (166a) and in (166b), a word final non-back vowel is followed by a clitic pronoun, but it surfaces as a glide in the sequences in (166a), whereas it is deleted in the sequences in (166b).

- (166) a. *disse-o* *0/[j] b. *sempre o disse* 0/^x[j]
 (he) said-it (he) always it said
 come-a *0/[j] *já ele a viu* 0/^x[j]
 (he) eats-it-FEM already he it-FEM saw
 cante-o *0/[j] *somente o espero* 0/^x[j]
 (he) sing-SUBJ-it (I) only it expect

The reference to the verbal category is also insufficient to account for the obligatory presence of [j], since in the sequences in (167) the glide does not surface, unlike in (166a).

- (167) *disse o João* 0/^x[j]
 said (the) John
 pede a prenda 0/^x[j]
 (he) asks (for) the-FEM gift
 cante o hino 0/^x[j]
 (he) sing-SUBJ the hymn

Besides these facts, there are also cases where both schwa deletion and [j] are possible. Such cases involve stressless function words, as shown in (168) (see chapter 5, section 4.1.2, for other phonological properties of clitic function words).

That the morphological category is not the determinant factor for the behavior found in (168) can be seen by the absence of a glide when the

schwa belongs to a stressed pronoun, as in (169), in contrast to (168a) above that includes stressless pronouns.

- (168) a. *não se ocUpou* 0/[j] (he) did not itself occupy
 já me aceitOU 0/[j] (he) has already me accepted
 sem te ouvIr 0/[j] without you hear
- b. *que ouvI* 0/[j] that (I) heard
 se a MarIa 0/[j] if the-FEM Mary
 de o ouvIr 0/[j] of it hear
- (169) *já Ele o disse* 0^x/[j] (he) already it said
 sem Ele a ver 0^x/[j] without he it-FEM see
 sem Este o sentir 0^x/[j] without the latter it feel

As we have seen in chapter 2, section 5, d'Andrade and Viana (1993) suggest that in order to account for this complex array of data the rule of schwa deletion must refer to the morphological category of the words involved: (i) in some cases it must refer to the category of the first word, as in the cases involving *grammatical* words in (168); and (ii) in other cases it must refer to the category of the second word, as in the forms with a clitic in (166a). However, the requirement in (i) is falsified by the data in (169), and the requirement in (ii) is falsified by the data in (166b). We believe such an analysis, in addition to be challenged by the data, fails to capture some of the most important factors that play a role in schwa deletion, that is, V2 stress, the prosodic structure of compounds and of clitic-host sequences, and the phonological behavior of clitics.

Anticipating some of the results presented in chapters 5, 6, and 7, we propose that schwa deletion applies to a prosodic word final vowel, as in (170) (on the feature specification of the segments that this rule targets, see note 146).¹⁴⁹

(170) Non-back Vowel Deletion (to be refined in chapter 6, section 3.1)¹⁵⁰

$$\left(\begin{array}{c} V \\ \text{-high} \\ \text{-back} \end{array} \right) \rightarrow \emptyset / [\dots _]_{\omega}$$

This explains why clitic words do not undergo this process, which is in general (nearly) obligatory.¹⁵¹ This formulation of the process also captures in an interesting way the asymmetry between verbal enclitics and proclitics illustrated in (166) above, since the former pattern like elements incorporated into the verb's prosodic word, but the latter do not, as we show in detail in chapter 5. As for compound-like structures, the absence of vowel deletion only when V2 is stressed and when the second word is the rightmost element of a compound-like structure suggests that these constructions are in fact compound prosodic words. This is the analysis proposed in chapter 6, where we show that non-back vowel deletion applies in prosodic word final position, but, similar to other vowel deletion processes (see sections 13, and 14), it is blocked if V2 bears the compound prosodic word stress.

The analysis we propose is thus capable of accounting for the facts described, and enables us to characterize this process as purely phonological, exceptionless, and operating postlexically.

As already pointed out above, the conditions on the realization of the glide are further discussed in chapters 5, 6 and 7.

13. Final round vowel deletion

As frequently reported in the literature, stressless underlying high round vowels in word final position may either surface as vowels, semivowels or be deleted, if followed by a word starting with a stressless vowel (e.g. Sá Nogueira 1938; Frota 1996, 2000: chap.2; Ellison and Viana 1996).

In this section, we will only concentrate on round vowel deletion which, when possible, is always an optional process.¹⁵² As the examples in (171) show, the quality of V2 seems to play no role in this process (cf. Ellison and Viana 1996).

(171) <i>bElo animAl</i>	0 [ɐ]	<i>saboroso almOço</i>	0 [a]
‘beautiful animal’		‘tasty lunch’	
<i>simpÁtico ouvINte</i>	0 [o]	<i>cArro usAdo</i>	0 [u]
‘nice listener’		‘old car’	
<i>gAto elegANte</i>	0 [ɛ]/[i]	<i>blcho herbÍvoro</i>	0 [ɛ]
‘elegant cat’		‘herbivorous animal’	

By contrast, V2 stress may affect round vowel deletion. In particular, as noted in Frota (2000) in certain configurations involving nearly adjacent word stresses, the presence of stress in V2 may block vowel deletion. In these cases, furthermore, the stress-level of the two words involved plays a crucial role. This is illustrated in the sentences in (172), taken from Frota (2000: 2.2.3.2) (in the examples, we include the relevant aspects of the prosodic phrasing of the sentences).¹⁵³

- (172) a. [o vestIdo ÂMbar]_φ foi vendido ontem *0
 the dress amber was sold yesterday
 ‘The amber dress was sold yesterday.’
- b. [o dançarIno]_φ [Ama]_φ a bailarina russa *0
 the dancer loves the dancer Russian
 ‘The dancer loves the Russian chorus girl.’
- c. [o bailarIno]_φ [ANda sempre]_φ de limusine preta 0
 the dancer drives always PREP limousine black
 ‘The dancer always drives a black limousine.’
- d. [o bailarIno]_I [ANtes de partir]_I falou com os amigos 0
 the dancer before PREP (to) leave talked with the friends
 ‘The dancer has talked with his friends before leaving.’

In (172a), vowel deletion is impossible because V2 bears word stress and ϕ -level stress. In (172b) the same happens, although the first word (W1) and the second word (W2) do not belong to the same ϕ , and both words bear ϕ -level stress. In (172c), by contrast, vowel deletion is possible because W2 does not bear ϕ -level stress, whereas W1 is the head of its ϕ -phrase. Finally, the sentence in (172d) shows that deletion is possible if W1 bears the intonational phrase stress, but W2 does not.¹⁵⁴

As for the sequences where W2 is not the head of its ϕ -phrase, and both W1 and W2 belong to the same phonological phrase, Frota (2000) considers that round vowel deletion is still impossible (in this case based on her own intuitions). We believe, however, that vowel deletion in this case is acceptable, at least for some speakers (including ourselves). The sentences in (173) illustrate these cases.

- (173) [o últImo ANjo negro]_φ revoltou-se contra Deus ?0
 the last angel black rebelled-himself against god
 ‘The last black angel rebelled against god.’

[as Oito Ânncoras pretas]_φ prenderam-se nos corais ?0
 the eight anchors black stuck-themselves in-the corals
 'The eight black anchors were stuck in the coral reef.'

Frota (2000) analyzes the cases where there is blocking of vowel deletion as a strategy available in EP to avoid sequences of adjacent stressed elements, or in other words to avoid stress clashes.¹⁵⁵ She further notes that clashes within- ϕ are more "tolerated" than clashes across- ϕ , which are strongly avoided if the rightmost clashing element is the head of the phrase. We believe the cases documented in (173) suggest an additional distinction in the within- ϕ situation: within a ϕ -phrase, clashes are more tolerated if the rightmost clashing element is not the head of the ϕ (as in 173) than if it is the head of this constituent (as in 172 above). As we will see in chapter 6, this greater tolerance of clashing stresses within ϕ is not found within compound structures, where vowel deletion in clashing contexts is totally forbidden if W2 is final in the compound structure. This behavior will then be related with the specificity of compound words' prosodization, as opposed to the regular concatenation of words within the ϕ constituent.

Although in her work only round final vowels are actually studied, Frota (2000) formalizes as a single process the deletion of final [u] and the deletion of final [ø]. Nevertheless, there are at least two differences between the two processes: (i) when followed by a [ø] or a [a], the first [ø] is not deleted but merges instead with the following vowel, and surfaces as [a]¹⁵⁶; (ii) when followed by a high vowel, the central vowel tends not to be deleted, and in some cases, [ø] merges with [u], yielding the low round vowel [ɔ] (see section 14 and chapter 5, section 4.3.1).

In fact, although central vowels and round vowels are deleted in similar contexts, it is not possible to merge the two in a single process. For example, if V1 deletion is seen as only possible if the two vowels are different, this implies that [u] is not deleted when followed by another [u], contrary to fact (see the relevant case in 171 above). In addition, if vowel deletion is restricted to apply only when V2 is non-high, the possibility of round vowel deletion when V2 is high is incorrectly ruled out.

We therefore reformulate Frota's (2000) account, as in (174), which includes the domain that the author defines to bound this process.¹⁵⁷ We further specify the round vowel as high in order for the rule not to apply to the exceptional cases where stressless low round vowels appear in prosodic

word final position, as in the final position of the first stem of morphological compounds or of clipped words (e.g. mIcr[ɔ]-ONdas ‘microwaves’ – see section 3).

$$(174) \begin{pmatrix} V \\ +\text{round} \\ +\text{high} \end{pmatrix} \rightarrow \emptyset / \left[\dots [\dots _]_{\omega} [V \dots]_{\omega} \dots \right]_I^{\max}$$

That the process applies across prosodic words within a compound intonational phrase is empirically supported. For example, the possibility of round vowel deletion in the example in (172d) shows that the presence of an intonational phrase boundary does not block vowel deletion. In such a case, round vowel deletion is possible, but only because the two I-phrases are grouped into a compound I-domain, since across non-compound I-phrases the rule is claimed not to apply.

Unlike the domain within which the process applies, the relevance of the prosodic word domain for its application is not discussed in Frota’s work. Nevertheless, in chapter 5 we will see that the formulation in (174) correctly predicts the behavior of round vowel deletion in sequences of words where either W1 or W2 is a clitic, and therefore does not form a prosodic word independent of its host: (i) the process does not apply between a verb and an enclitic, because the enclitic is incorporated into the verb’s prosodic word; (ii) it may apply between an enclitic and a following word, because the enclitic becomes the final element of the prosodic word it incorporates into; (iii) it does not apply between a proclitic and its host, since the proclitic is not an independent prosodic word; and (iv) it may apply between a word and a following proclitic, because the proclitic is prosodically attached to the following prosodic word (see chapter 5 for the details).

Finally, the forms in (175) illustrate the non-application of round vowel deletion *within* a prosodic word.

(175) <i>tÁbua</i>	*0	<i>actuaçÃo</i>	*0
‘board’		‘action’	
<i>duodEno</i>	*0	<i>coibIr</i>	*0
‘duodenum’		‘(to) restrain’	

The optional nature of this process, as well the lack of exceptions and the reference to pure phonological information only, allow us to classify it as a postlexical phenomenon.

14. Final central vowel deletion

Stressless central vowels in word final position may also be deleted, if the following word starts with a stressless vowel different from [ɐ]. This is illustrated with the forms in (176), taken from Sá Nogueira (1938).

- (176) *Ela errAva* 0 ‘she had made mistakes’
 a filha orAva 0 ‘the daughter had prayed’

Sá Nogueira reports that, in addition to the vowels represented in (176), the sequences where V2 is [u] also allow for central vowel deletion. In our dialect, however, central vowel deletion in this context is a marked possibility (see 177).

- (177) *a vElha usAva* ^x0 ‘the old-woman used’
 uma cAsa usAda ^x0 ‘an old house’

Moreover, and also unlike Sá Nogueira (1938), we establish a contrast between central vowel deletion with [i], as in (178), and central vowel deletion with non-high vowels. Only in the latter case do we believe vowel deletion regularly applies.

- (178) *a rAra imAgem* ^x0 ‘the rare image’
 uma lINda idOsa ^x0 ‘a beautiful old-lady’

We should note, nevertheless, that Ellison and Viana (1996) follow Sá Nogueira’s judgments. They claim, additionally, that when V2 is a high vowel either the central vowel is deleted or V2 semivocalizes. Since we also find the latter possibility a marked one, we believe that there may be some variation in the way speakers treat these vowel sequences.¹⁵⁸

With respect to stress-clash situations, central vowels appear to display a behavior similar, though not identical, to round vowels, as shown in (179).

- (179) a. [trINta Éguas puras]_φ foram vendidas ontem ?0
 thirty mares pure were sold yesterday
 'Thirty pure mares were sold yesterday.'
- b. [a vElha Obra]_φ foi acabada ontem *0
 'The old work was finished yesterday.'
- c. [a dançarIna]_φ [OUve]_φ muita música ?0
 the dancer hears lots music
 'The dancer hears lots of music.'
- d. [a bailarIna]_φ [Olha sempre]_φ para rapazes bonitos 0
 the dancer stares always at boys pretty
 'The dancer always stares at pretty boys.'
- e. [a bailarIna]_I [hOje reformada]_I falou com os amigos ^x0
 the dancer today retired spoke to the friends
 'The dancer – nowadays retired – spoke to their friends.'

These facts therefore suggest that deletion is possible when W1, but not W2, is the head of its ϕ (cf. 179d). By contrast, if W1 and W2 belong to the same ϕ and W2 is the head of this constituent, central vowel deletion is particularly disfavored (cf. 179b). When V2 bears word stress and is not the head of ϕ and when both W1 and W2 are the head of their ϕ 's, deletion is possibly a marked option (cf. 179a and 179c, respectively).

As for the marginal status of central vowel deletion in (179e), we believe it is not due to the presence of a clash situation, but rather to the fact that this process, unlike round vowel deletion, is not operative across intonational phrases, even if they are included within a compound intonational domain. An argument in favor of this view comes from the fact that the same tendency not to delete the central vowel in the same prosodic environment is found if V2 is unstressed, as in (180).

- (180) [A bailarIna]_I [organizAda como sempre]_I falou com os amigos ^x0
 the dancer organized as usual spoke to the friends
 'The dancer – always very organized – spoke to their friends.'

Although this description requires further investigation, it certainly suggests that this process applies within a non-compound intonational phrase.

As we will see in chapter 6, the deletion of the central vowel is obligatorily blocked within compound structures if V2 bears the compound stress, regardless of ϕ -level prominence. In this respect, central vowel dele-

tion will thus be shown to pattern exactly like round vowel deletion, and also like non-back vowel deletion.

The fact that central vowel deletion is impossible between a verb and an enclitic, while it is possible between a word and a proclitic, as we will see in chapter 5, section 4.3, suggests that it applies at the right-edge of the prosodic word domain (see also section 4.3 of chapter 5 for the details on the prosodization of clitics). The forms in (181) further show that central vowel deletion does not apply within the word.

- (181) *caÓtico* *0 *extraordinÁrio* *0
 ‘chaotic’ ‘extraordinary’
 aeropOrto *0 *aeronÁutica* *0
 ‘airport’ ‘aeronautics’

We therefore propose that this process optionally applies at the right edge of the prosodic word, within the intonational phrase domain, as in (182).

- (182)
$$\left[\begin{array}{c} V \\ -\text{round} \\ +\text{back} \end{array} \right] \rightarrow \emptyset / \left[\dots [\dots]_{\omega} \left[\begin{array}{c} V \\ -\text{high} \\ \alpha\text{back} \\ \alpha\text{round} \end{array} \right] \dots \right]_{\omega} \dots]_I$$

Central vowel deletion has the properties of a postlexical phenomenon: (i) it is general; (ii) it is optional; and (iii) it refers only to phonological information.

15. Syllable degemination

Syllable degemination is a process whereby, in a sequence of two similar syllables – whose precise make-up will be described below – the first one is deleted (e.g. Sá Nogueira 1938; Villalva 1994; Frota 1997, 2000).

Syllable degemination (SD) is not a general process word internally (cf. Villalva 1994; Frota 2000). Although in derived environments there are some forms that seem to be the result of the application of SD (cf. Sá Nogueira 1938; Villalva 1994), this is not a productive process of EP: while in the forms in (183a) syllable degemination is strictly obligatory, in forms in (183b) it is totally excluded.

- (183) a. **bondadOso* / *bondOso* 'kind' (cf. *bondade* 'kindness')
**caridadOso* / *caridOso* 'charitable' (cf. *caridade* 'charity')
**idadOso* / *idOso* 'aged' (cf. *idade* 'age')
- b. *cuidadOso* / **cuidOso* 'careful'
predadOr / **predOr* 'predator'
nadadOr / **nadOr* 'swimmer'

This is further supported by the examples in (184), where similar adjacent syllables do not trigger syllable degemination.

- (184) *bebedEira* [bibi]/*[bi] *tutoriAl* [tutu]/*[tu]
'drunkness' 'tutorial'
cacarejAr [kəkə]/*[kə] *memorizAr* [mimu]/*[mu]
'(to) cackle' '(to) memorize'

The same is not true, however, in larger prosodic domains. Frota (2000) shows that the process applies optionally across word boundaries within the intonational phrase domain, if two adjacent syllables have some degree of similarity, and the first one is unstressed, as in (185).

- (185) [o meu pai herdou [um cAMpo poluído]_φ]_I [pupu]/[pu]
the my father inherited a field polluted
'My father inherited a polluted field.'
[[o cAMpo]_φ [podIa estender-se]_φ até ao ribeiro]_I [pupu]/[pu]
the field could go-itself up to-the river
'The field could go up to the river.'

Despite the fact that the degree of similarity between the two syllables is not systematically investigated in Frota (2000), the author suggests that the vowels involved should have the same height, and the more similar the syllables are the more likely the process is to apply. The examples in (186) show, however, that the process occurs only with syllables ending in [u] and [i], but not with syllables ending in [ə] or [ɪ].¹⁵⁹

- (186) a. *Ele bEbe beblDas espirituosas.* [bibi]/[bi]
he drinks drinks alcoholic
'He drinks alcoholic drinks.'

116 *Phonological phenomena*

<i>Ele tEme motINs desnecessários.</i>	[mĩmu]/[mu]
he fears insurrections unnecessary	
‘He fears unnecessary insurrections.’	
<i>um gAto tolInho</i>	[tutu]/[tu]
a cat silly	
‘a silly cat’	
<i>um bocAdo dorIdo</i>	[dudu]/[du]
‘a bit hurt’	
b. <i>transpOrta tapEtes persas</i>	[tɛtɛ]/*[tɛ]
‘(he) carries Persian carpets’	
<i>lIMpa passEios</i>	[pɐpɐ]/*[pɐ]
‘(he) cleans walks’	
<i>Dĩli libertAda</i>	[lili]/*[li]
‘Dili free’	
<i>GÂNdi dilacerAdo</i>	[didi]/*[di]
‘Gandhi dilacerated’	

Notice that syllable degemination applies when open syllables are involved, but not when the two syllables are closed by a consonant, even if the syllables are identical, as shown in (187).

(187) <i>pEdes descUlpas</i>	[di]di]/*[di]	‘(you) ask forgiveness’
<i>ignÓbil biltrAgem</i>	[bi]bi]/*[bi]	‘abject dishonesty’
<i>fÊmur murdIdo</i>	[murmur]/*[mur]	‘femur bitten’

In addition, the first syllable cannot be deleted if both syllables have a branching onset, as shown in (188).

(188) a. <i>Não cOMpre pretInho.</i>	[pri]pri]/*[pri]
not buy-SUBJ black-DIM	
‘Do not buy (it) black.’	
b. <i>COMpro prodUtos biológicos.</i>	[pru]pru]/*[pru]
(I) buy products biologic	
‘I buy biologic products.’	
c. <i>Era um ciclo clubIsta.</i>	[klu]klu]/*[klu]
(it) was a cycle clubistic	
‘It was a clubistic cycle.’	

Finally, syllable degemination seems nevertheless possible, if the second syllable, but not the first one, has a filled coda, as in (189).

- (189) a. *quisEste terminAr* [tʰitʰr]/[tʰr] b. *pOdes detEr* [dizdi]/*[di]
 ‘(you) wanted (to) finish’ ‘(you) can stop’
 um gAto toscAno [tutuʃ]/[tuʃ] *gAtos tollInhos* [tuʃtu]/*[tu]
 a cat Tuscan cats silly
 ‘a Tuscan cat’ ‘silly cats’

The conditions on syllable degemination that we believe tend to be observed are summarized in (190).¹⁶⁰

(190) Conditions on syllable degemination

- a. The rhyme of the first syllable is non-branching
- b. The onset of both syllables is non-branching
- c. The onset of the two syllables is filled by the same consonant
- d. The nucleus of the two syllables is filled by a back high vowel

Frota (2000) assumes that this is a prosodic word domain limit rule, and she formulates the rule in two steps: (i) vowel deletion, which may apply across an intonational phrase boundary, within a compound I-domain; and (ii) consonant deletion, which is bound by a non-compound intonational phrase domain (cf. Frota 2000: 72). We reformulate her proposal as in (191), in order to restrict the process to apply only when the conditions stated above are met.

- (191) a. $V_1 \rightarrow \emptyset / [\dots [\dots C_a _]_{\omega} [C_a V_2 \dots]_{\omega} \dots]_I^{\max}$
 b. $C_a \rightarrow \emptyset / [\dots [\dots _]_{\omega} [C_a \dots]_{\omega} \dots]_I$

(where the syllables headed by V_1 and V_2 conform to the conditions stated in 190)

Note that, besides intuition, a clear indication that syllable degemination in EP deletes the first rather than the second of two similar syllables is provided by the cases where the two relevant syllables are not identical. In such cases, the shape of the syllable that remains in the string corresponds to the second rather than the first one (e.g. *prometo terminar já* [prumétutirminár]/[prumétirminár] ‘(I) promise to finish shortly’).

Apart from the fact that the rule applies optionally between prosodic words that are directly attached to the phonological phrase, as in the examples in (185) above, the data considered in chapter 5 will also corroborate the status of this rule as a prosodic word limit rule. In particular, syllable degemination will be shown not to apply to the final syllable of the verb if it is followed by an enclitic. This is consistent with the description of the process as a prosodic word domain limit rule, and with the analysis proposed in chapter 5 that enclitics incorporate into the host's prosodic word. Conversely, the process may apply between a word and a proclitic, which is analyzed as attached to the following prosodic word, and thus provides the context for the rule to apply.

Finally, this process is optional and applies across words without exceptions, under the phonological conditions expressed in (190).¹⁶¹ It should therefore be considered a postlexical process.

16. Initial stress

On the basis of the observation of syllable durations within the word, d'Andrade and Viana (1989) argue that, besides primary word stress, words also show echo stresses, which are computed with reference to the primary stress position. This "rhythmic wave" consists of alternating strong-weak beats that proceed from the primary word stress leftwards.

As noticed in d'Andrade and Viana (1999) and Frota and Vigário (2000), however, the results of d'Andrade and Viana (1989) only allow for a clear distinction between three types of syllables: the one that bears main stress, the initial syllable that bears initial stress, and all the remaining syllables within the word. D'Andrade and Viana (1999: 88) suggest, furthermore, that only primary stresses and initial stresses are likely to be found in the speech continuum, while echo stresses are usually not detectable. These suggestions are confirmed by the results of Frota and Vigário (2000), where, besides the syllables bearing word stress, the syllables of uttered sentences reported to be perceived as prominent always correspond to the initial syllables of a (prosodic) word (see also chapter 5, section 4.3.2).¹⁶² In other words, in this study no word internal prominent syllables were found.¹⁶³

Mateus and d'Andrade (2000: 6.5) draw a clear separation between the two types of secondary prominence: echo stresses are claimed to be rhythmic, whereas the initial stress is argued to be non-rhythmic and positionally determined.

Two important questions are relevant for us here: (i) what is the domain of the positionally determined secondary prominence?; (ii) when is this prominence assigned? These questions are related, as we will see in the next paragraphs.

There are several arguments for the postlexical status of the initial prominence. First, it does not refer to morphological or lexical information, unlike word stress. Second, it is optional, as suggested in d'Andrade and Viana (1999) and empirically shown in Frota and Vigário (2000). Third, it includes proclitics in its domain, as we will see in detail in chapter 5 (see also d'Andrade and Viana 1999 and Frota and Vigário 2000). Fourth, it is more frequent in I-initial position, as empirically shown in Frota and Vigário (2000). Since prosodic constituents higher than the prosodic word are not yet present in the lexical component (cf. chapter 1, section 3.2), and given that proclitics seem to attach to their hosts only postlexically (cf. chapter 4), all these facts strongly support the postlexical nature of initial stress assignment.

The separation of rhythmic stress and word stress computation is also claimed in Roca (1986), who assigns non-primary stress in languages like Spanish and Italian to the postlexical component. Van der Hulst (1996) further notes that non-primary accents refer to phrasal and larger domains and display properties that are used as diagnostics for postlexical rules, such as optionality and freedom of exceptions (see also Roca 1986; Nespor 1999b). The assignment of rhythmic stress to the postlexical component is also claimed in Pereira (1999) for EP. Although initial prominence assignment in EP cannot be treated as (other) rhythmic stresses, since it is positionally determined, it certainly contributes for the rhythmic impression of EP, and it clearly behaves as a postlexical phenomenon.

Finally, the fact that proclitics are included in the domain of initial stress assignment suggests in addition that the relevant notion of word is phonological rather than morphological.

Within the spirit of the end rule presented in d'Andrade and Viana (1999), we formulate initial stress assignment in (192).¹⁶⁴

(192) Initial stress assignment (optional)

The first syllable of a prosodic word is a peak.

In chapter 5, section 4.3.2, we present illustrative data on initial stress and discuss further its domain of occurrence.

17. Emphatic stress

Although used very frequently, in particular in politicians' speech, or TV news reports, emphatic stress is not commonly described in EP literature. As illustrated in (193), this stress may be assigned to the initial syllable of a word.¹⁶⁵

(193) *Foi DEterminANte a actuação do comandante.*

(it) was crucial the action of-the commandant
'The action of the commandant was crucial.'

Exige-se a INstauração de um inquérito.

demand-they the opening of an inquiry
'They demand the opening of an inquiry.'

Isto é SUbErbo!

'This is great!'

We should notice that what we mean by emphatic stress does not coincide entirely with other uses of the term, namely, by Moraes Barbosa (1965: 8.9). For Moraes Barbosa, emphatic stress may be assigned to any syllable of a word. In our view, however, this free distribution of stress is only possible in particular pragmatic conditions: namely, in a situation where the phonetic form of a previous production is misperceived and thus must be clarified, as in (194).¹⁶⁶

(194) *Eu disse preferIr e não preterIr.*

'I said *prefer* and not *reject*.'

These particular cases are distinct in distribution and meaning from the phenomenon we are describing here. As pointed out in Jackendoff (1972: 242) for English examples similar to (194), these cases involve a phonological but not a semantic contrast. In the case of emphatic stress, however, no contrast is established, as this particular type of stress is used with the function of highlighting the relevant word in any discourse context.

Another type of stress we should distinguish emphatic stress from is the Focal stress, described in section 19 below. As we will see then, some of the properties of phonological Focus are the following: (i) the presence of focus has implication for the semantic reading of the sentences; (ii) the presence of focal stress affects the default prominence pattern at the level of the intonational phrase; (iii) only *content* words seem to be able to be fo-

cused; (iv) the focal stress falls on the syllable that bears primary word stress; (v) a specific pitch accent is associated with the stressed syllable of the focused element; and (vi) the tonal space after the focused element becomes narrower (cf. Frota 1994, 2000 and Vigário 1997a, 1998a for the phonological details of phonological focus in EP). Although, to our knowledge there are no other studies characterizing emphatic stress in EP, it seems clear that all these properties are absent in the realization of emphatic stress, and thus they clearly distinguish phonological Focus from emphatic stress (as the examples in 193 partially illustrate).¹⁶⁷

There are a number of facts that point to the postlexical nature of emphatic stress assignment: (i) it is not sensitive to morphological or lexical information; (ii) it is optional; (iii) it never blocks vowel reduction; and (iv) it may be assigned to elements that become prosodic word initial only postlexically, namely, proclitics (see chapters 4 and 5).

Finally, the fact that this stress can occur in initial syllables of unprefixated words, in initial syllables that belong to prefixes, as well as in proclitics (as it will be shown in chapter 5) suggests that the relevant concept here is the phonological rather than the morphological word.

Taking into consideration the generalizations described above, we propose to account for emphatic stress by (195).¹⁶⁸

(195) Emphatic stress assignment (optional)

Emphatic stress is assigned to the first syllable of a prosodic word.

This type of stress is further discussed in chapter 5, section 4.3.2 (see also Appendix I for collected data).

18. Tonal association

Non-boundary tonal events are commonly described to be associated with stressed syllables in EP (cf. Viana 1987; Frota 1994–2002b; Falé 1995; Vigário 1998a). In the literature on EP intonation that we are aware of, elements that can be regarded as clitics are not assigned pitch accents. Thus, the presence of a pitch accent may be seen to cue the stressed status of a given word. The stylized contour in Figure 1 illustrates the association of three pitch accents with three stressed positions in a simple declarative sentence.¹⁶⁹

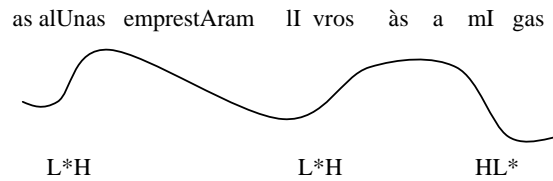
122 *Phonological phenomena*

Figure 1. Stylized contour and pitch accent transcription of sentence (1) of Vigário (1998a: 191) *As alunas emprestaram livros às amigas* ('The students lend some books to their friends.').

In agreement with the observations found in the literature, Frota (2000) claims that the minimal tune of a sentence in EP consists of a pitch accent associated to the head of the intonational phrase plus a final boundary tone. Pitch accents may also be found in sentence earlier positions, which correspond, in general, to the head of a ϕ -phrase.

There are, nonetheless, two situations where tonal information may be found associated to a prosodic word that does not correspond to the head of either I or ϕ . A pitch accent may appear inside a non-initial ϕ -phrase, associated to a non-head prosodic word if the head is also assigned a pitch accent. This is illustrated in (196a), taken from Frota (2000: 4.2.2), and in (196b), based on Vigário's (1998a) data.¹⁷⁰

- (196) a. *uma subida progressiva* 'a progressive rise'
 (i) --- PA
 (ii) PA PA
 (iii) PA --- (not attested)
 b. *somente telas* 'only canvas'
 (i) --- PA
 (ii) PA PA
 (iii) PA --- (not attested)

The second situation where a tone may be associated with a non-head prosodic word is in intonational phrase initial position (cf. Vigário 1998a: 113). As shown in different works on EP intonation, an H tone may be associated with the final syllable of the intonational phrase initial constituent.¹⁷¹ The fact that the initial constituent of most sentences found in the literature often corresponds to a non-branching ϕ has led to the assumption that the initial H is, like other pitch accents, assigned to the head of a ϕ -domain (e.g. Frota 2000). However, sentences where an initial ϕ contains

two prosodic words, such as the one in (197) (see the F0 contour 59 of Vigário 1998a: 222), show that the initial H may indeed be associated with a prosodic word that does not correspond to the head of ϕ .¹⁷²

(197) [[[Apenas]_ω [as [professoras]_ω] _φ ...]_I ‘only the teachers...’
 |
 H

Intonational information may, thus, provide evidence for the prosodic word domain.

Like other Romance languages, EP has no lexical tones and thus intonational phenomena may be straightforwardly characterized as postlexically assigned.

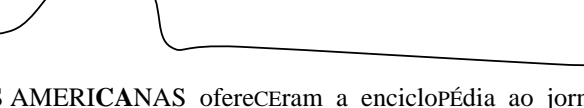
19. Focal stress

Focal stress in EP was extensively studied by Frota (1994b, 1997, 2000), who analyzes affirmative sentences, and, to a less extent, by Vigário (1997, 1998a), who analyzes negative sentences.¹⁷³

From a semantic point of view, focal stress identifies narrow/contrastive focus, as exemplified in (198a) and (198b) (from Frota 2000: 97 and Vigário 1998a: 54, respectively).¹⁷⁴ In the two examples, a contrastive reading is preceded by a context sentence.¹⁷⁵

- (198) a. *Foram as africanas que ofereceram a enciclopédia ao jornalista.*
 'It were the African girls that gave the encyclopaedia to the journalist.'
(Não.) As AMERICANAS ofereceram a enciclopédia ao jornalista.
 (No.) the American (girls) gave the encyclopaedia to-the journalist
 '(No.) The American girls have given the encyclopaedia to the journalist.'
- b. *As garotas deram livros às velhotas.*
 'The girls have given some books to the old ladies.'
(Não.) As garotas EMPRESTARAM livros às velhotas.
 (No.) the girls lent books to-the old-ladies
 '(No.) The girls have lent some books to the old ladies.'


a.



AS AMERICANAS ofereCEram a enciclopÉdia ao jornalISTA

H^*+L

b.



As americANas ofereCEram a enciclopÉdia ao jornalISTA

H^*

Like other EP intonational and prominence phenomena (to the exception of word stress), focal stress has the properties of a postlexical phenomenon.

20. Summary

In this chapter, we have surveyed a number of phonological phenomena of EP. We have argued for the separation of the phenomena that refer to morphological information and that may have exceptions from the phenomena that refer only to phonological information. The first kind of phenomena was assumed to apply within the lexical component, whereas the latter kind was claimed to either (i) be active in the lexical component, or (ii) operate

postlexically. The *lexical* phenomena described in this chapter were characterized to apply before words are combined, so that any change in the context caused by the concatenation of words never affects this type of operations. We have seen, in addition, that the occurrence of *postlexical* phonological phenomena never affects (triggers or blocks) processes that are argued to operate in the lexical component. By contrast, postlexical phonological phenomena were shown to operate automatically in environments that may be obtained through the postlexical concatenation of words.

Table 1 shows a classification of the phenomena claimed in this chapter to be either lexical or postlexical.

Table 1. Classification of the phonological phenomena described in this chapter, according to their locus of operation and to the relevance of the prosodic word domain for their occurrence.¹⁷⁹

<i>Lexical phonological phenomena</i>	<i>Reference to the Prosodic Word</i>
Word stress	NO
Vowel reduction	NO
Theme vowel deletion	NO
Final nasal diphthongization	NO
Heterosyllabic /e/-centralization	YES
Glide insertion to break a hiatus	YES
Lowering of stressless vowels in final syllables closed by /r/	YES
Initial /r/-strengthening	YES
Initial vowel realization (neutralization)	YES
<i>Postlexical phonological phenomena</i>	
Initial vowel realization (feature specification)	YES
V2 semivocalization	YES
Final non-back vowel deletion	YES
Final round vowel deletion	YES
Final central vowel deletion	YES
Syllable degemination	YES
Initial stress	YES
Emphatic stress	YES
Tonal association	YES
Focal stress	YES

It identifies, furthermore, the phenomena argued to apply with reference to the prosodic word domain. Further evidence in favor of the relevance of the prosodic word domain for the application of a given process is found in chapters 5 and 6.

In the following chapters we make extensive use of these findings. It is our contention that (i) except for precompiled phrasal rules (cf. Hayes 1990 and chapter 2, section 3.1 of this book), lexical phenomena do not affect elements that are combined postlexically; and (ii) only postlexical (pure) phonological phenomena constitute reliable evidence for the prosodization of elements that are combined postlexically. In accordance with (i), we propose in chapter 4 that the phenomena classified here as lexical do not apply between the verb and the clitic pronoun because this sequence is obtained postlexically. These facts are thus seen to argue against the view that clitic pronouns in EP are inflectional affixes. In accordance with (ii), in chapters 5 and 6 we assess the way clitics and other postlexically combined words are prosodized in EP resorting only to the phenomena shown in this chapter to apply postlexically.

Personal copy

⁷⁸ We are using the traditional articulatory classification of EP segments (the terms palatal and velar are also used instead of front and back, respectively). In the standard variety of EP, only the back vowels [u, o, ɔ] are rounded. For a classification in terms of distinctive features, as defined in Chomsky and Halle (1968), see, for example, Mateus (1975: 72).

⁷⁹ For extensive discussion and analysis on the data described here see, in particular, d'Andrade (1988, 1997), d'Andrade and Laks (1992), Pereira (1999) and Mateus and d'Andrade (2000: chap.6). In earlier accounts, stress is computed from right to left, constituents are binary, and left-headed, and the head of the rightmost constituent is assigned the primary stress. In all of these analyses, some vowels are exceptionally marked in a given lexical entry, so that they are not assigned stress. In Mateus and d'Andrade (2000) phonetically null class-markers are also assigned a rhythmic position. Within the verb system, most tense suffixes are usually claimed to be lexically marked in order not to bear word stress.

⁸⁰ In the examples, whenever relevant we mark the vowel that bears word stress with a capital letter. For ease of exposition, we use in general EP orthographic system, and only when justified we employ the International Phonetic Alphabet, inside square brackets or slashes, to indicate phonetic or phonological forms, respectively.

⁸¹ There are also triplets of words that have the same underlying segmental structure and differ with respect to stress placement, as *sabiÁ* (a kind of bird), *sabla* '(he) knew', *sÁbia* 'wise woman'. In this case, the words belong to different morphological classes: respectively, nominal, verbal, and adjectival classes. As pointed out by Mateus and d'Andrade (2000: 6.3), this sort of data argues against the view that EP stress is quantity sensitive (as claimed, for example, by Brandão de Carvalho 1989).

⁸² In this chapter we use the following abbreviations: PAST: indicative past perfect; IMP: indicative past imperfect; SUBJ: subjunctive present; 1/2/3Sg: 1st/2nd/3rd/person singular; 1/2/3Pl: 1st/2nd/3rd person plural; MASC: masculine; FEM: feminine; PL: plural; DIM: diminutive. Notice that in some of the examples in (4) the theme vowel does not surface due to the application of a rule specific of the verb system that deletes the theme vowel when it is followed by another vowel (see section 3.3).

⁸³ We should add that enclitics behave like post-tonic syllables in that they are never assigned secondary stress in EP. This is already observed in Gonçalves Viana (1883), who notes the contrast between (some dialects of) Spanish and EP with respect to posttonic secondary stress assignment, as illustrated in (i), taken from Gonçalves Viana (1883: 62) (boldface signals secondary stress).

- | | | |
|----------------|---------|---------|
| (i) PORTUGUESE | SPANISH | |
| tÚmulo | tUmulo | 'grave' |

dizIa-lhe	dicIale	‘(he) used to tell him’
-----------	---------	-------------------------

⁸⁴ “Vowel reduction” is the traditional term used in EP literature for the phenomenon under observation in this section. We should notice that schwa, traditionally transcribed as [ə] in the literature on EP, is a very reduced central vowel. In the words of Gonçalves Viana (1883: 4): “[c]’est un *e* muet, comme on l’appelle généralement, bien plus étouffée, bien plus fermé, cependant, que l’*e* français de *me, le*.”

⁸⁵ Brandão de Carvalho (1989, 1994) considers VR a regular phenomenon, which applies to *light* syllables. As for the vocalic nuclei in open syllables that fail to undergo VR, they are claimed to be *heavy*. The examples in (12) seem, nevertheless, to pose a problem for the assumption that VR is a general phenomenon, since the same underlying vowel, in the same syllabic environment, undergoes VR in some word-forms and surfaces unreduced in other (related) word-forms. In addition to this, under this approach the fact that syllables closed by /s/ or /t/ in non-final position do not prevent vowel reduction implies the postulation that such syllables are light, unlike those closed by /l/. Finally, there seems to be no other evidence for a distinction among vowels based on weight in EP. To be more accurate, word stress assignment is argued to be sensitive to weight by some authors (e.g. Brandão de Carvalho 1989, followed in Morales-Front and Holt 1997). However, Mateus and d’Andrade (2000: 6.3) and Pereira (1999) present a number of arguments against this view (see also section 3.1).

⁸⁶ The rule that centralizes /i/ formalized in Mateus (1975: 26), d’Andrade (1977: 46), Andrade (1980: 48) includes a(n optional) final consonant. This formulation is meant to account for the final [i] in verb forms such as *pArtes* ‘(you) leave’, *Abres* ‘(you) open’. Again, this centralization seems to occur only in verbs, for in the noun system there are several instances of final unstressed [i] followed by a consonant (cf. i).

(i) lÁpis ‘pencil’	Útil ‘useful’	mÁrtir ‘martyr’
oÁsis ‘oasis’	fÁcil ‘easy’	
grÁtis ‘free’	diffícil ‘hard’	

It should be noticed that in early generative accounts some of the words presented in (i), namely those superficially ending in liquid, are assumed to have an underlying final /e/ (cf. d’Andrade 1977; Andrade 1980). Observing that the centralization of stressless final /i/ must follow the deletion of final stressless /e/, which is argued to apply very early in the derivation, Andrade (1980: 49) concludes that the centralization that affects /i/ must be specific to the verbal system.

⁸⁷ It is often mentioned in the literature that words ending in stressless [i] are, nevertheless, rare in EP. We believe this is not because of vowel quality facts, but rather results from the exceptional stress pattern that these words display. Since such words include no overt class marker, and given that stress in the nominal

system usually falls on the last vowel of the stem (see section 3.1), these forms must be lexically marked in order for the stress to fall on the penultimate syllable, unlike in words such as *colibrí* 'id.', *sagui* 'tamarin', *javali* 'wild pig'. Finally, we should point out that there are also a number of stressed prefixes ending in a stressless /i/ where centralization does not apply (cf. section 6.2.2).

⁸⁸ Diphthongs formed by a central vowel and [w] may constitute an exception to this generalization. For example, *saudar*, may be realized as [səwdár], although the most common realization of this word is [sodár]. As for the latter possibility, the most straightforward analysis is that an assimilation process has applied, changing [aw] to [ow] (the same type of assimilation is observable, for example, in the verbal system—see note 20). Thus, in the latter case VR does not apply because [o] is part of a diphthong. The realization of [ow] as [o], in turn, is a more general process of the Lisbon variety of EP (e.g. Mateus 1975: 42). The assimilation analysis may also explain the typical realizations of words like *saudade* 'nostalgia' and *saudoso* 'longing' with a [o].

⁸⁹ The regular realization of these vowels as low is also referred in section 3.7.

⁹⁰ The realization of initial vowels is discussed in section 3.9.

⁹¹ This description excludes stressless vowels followed by a fricative in word final position, since they surface as reduced vowels, as shown in (i).

- (i) ourIves [i]/0 ANtes [i]/0 alfEres [i]/0
 'goldsmith' 'before' 'second lieutenant'

We discuss the obligatory low realization of vowels in word final syllables closed by sonorants in section 3.7.

⁹² Only the compound linking vowel (–o) is usually reported to escape VR (e.g. Villalva 1994). The forms in (21) suggest, however, that this is a more general behavior of non-high vowels (–o or –e) in non-final stems of morphological compounds (see also section 6.2.3).

⁹³ See sections 6.3 on clipping in EP. Notice that the adverb *inclusíve* 'comprehensively', which coexists with *inclusivamente* and is realized by many speakers with a low final vowel may also be analyzed as an instance of clipping. We should further point out that, to our knowledge, the generalization described in (vii) was not previously noticed in the literature.

⁹⁴ Mateus (1997a) assigns vowel reduction to the postlexical component. The argument for that claim is based on the behavior of unstressed vowels in future and conditional tenses: when a clitic pronoun is inserted between two morpheme boundaries, stress is assigned as if the verb form contains two words and the stressed vowels do not undergo vowel reduction (e.g. *falar-te-Ei*, [fəlártjɛj] '(I) will speak to you'); by contrast, when no clitic occurs, vowel reduction applies regularly to unstressed vowels (e.g. *falarEi*, [fələɾɛj] '(I) will speak'). This view presupposes that the verb forms inserted in the future and conditional are underlyingly identical whether the clitic is present or not. In chapter 4, we show

that the two cases are obtained through different mechanisms, which explain the different behavior of the relevant vowels with respect to vowel reduction, while maintaining the assumption that this is a lexical phenomenon.

⁹⁵ We assume the underlying form of the 1st person singular suffix to be /u/ rather than /o/ because this vowel always surfaces as high, and there is no evidence for its mid status (cf. also d'Andrade 1977: chap.4). The same applies for the underlying form of the nominal class marker *-o* (see Andrade 1980: note 37 for a similar remark).

⁹⁶ In the case of 1st and 3rd person singular of the perfect indicative the theme vowel is not deleted possibly because it is followed by a glide: in the 1st conjugation, the theme vowel /a/ undergoes a specific process of assimilation under the influence of the glide, both in the 1st and in the 3rd person singular; in the 2nd and 3rd conjugations the theme vowel and the following glide undergo total assimilation in the 1st person singular, and thus a [i] surfaces; in the 3rd person singular, the theme vowel is not deleted also because it is followed by a glide. This is shown in (i) (for a similar analysis see e.g. Mateus 1983: note 18).

(i) falei < /fal a + j/	bati < /bat e + j/	parti < /part i + j/
'(I) spoke'	'(I) hit'	'(I) left'
falou < /fal a + w/	bateu < /bat e + w/	partiu < /part i + w/
'(he) spoke'	'(he) hit'	'(he) left'

We should notice, in addition, that fusion of a stressed high vowel and a palatal glide also occurs in non-verbal environments, as in *funis* 'funnel' (< funijs < /funil+s/) (cf. Morales-Front and Holt 1997), and the special case of assimilation involving the sequence /aj/ has also happened in other periods of the language in words such as *leite* (lat. *lacte*), and in the suffix *-eiro*, which is etymologically related to the suffix *-ário*. In Mateus and d'Andrade (2000: 4.2.2.1.2) a slightly different proposal is made, since these authors assume that EP has no underlying glides. Theme vowel deletion is claimed to apply only in the present tenses and thus the fact that the theme vowel is followed by a suffix starting with a vowel in the past tenses does not trigger vowel deletion.

⁹⁷ The theme is defined as the verb stem plus the theme vowel.

⁹⁸ We believe, however, that the nasalization of the vowel is possible in at least some of the cases where the prefix *en-* is followed by a nasal consonant: *en+madeirado* 'in between wood'; *en+muralhado* 'in between walls'. It is possible that speakers show some variability in these cases.

⁹⁹ This analysis, proposed in Morales-Front and Holt (1997), suggests that the final fricative is syllabified in coda position in the lexical component. Thus, the glide surfaces as non-back, even though the fricative may end up as the onset of the following syllable postlexically. The same is true for words ending in /l/ that undergo pluralization. Notice, nevertheless, that coda fricatives need not to be fully specified in the lexicon, since, if these consonants are simply specified as [-back],

the right result with respect to the palatal glide is obtained. See Mateus and d'Andrade (2000: 2.4.2) on the underspecification of coda consonants.

¹⁰⁰ In the case of words ending in stressed nasal vowel, as *fim* 'end' and *atum* 'tuna', we are assuming that the final glide, which has the same value for the features [back] and [high] as the preceding vowel, fuses with that vowel (the same type of process seems operative in the verbal system—see note 20). In forms like *patrão/patrões* 'boss(es)', also morphologically related to words like *patronal* 'relative to bosses', we assume that the underlying form of the stem is /patro[+nas]/ and the /o/ dissimilates to [ɐ] when followed by the nasal glide [w̃]. Notice that oral diphthongs of the form [ow] do not surface as such in the Lisbon variety of EP (e.g. Mateus 1975: 42) and there are no surface nasal diphthongs of this form either.

¹⁰¹ Some forms also exist where nasality is absent in etymologically related forms: e.g. *jejum/jejuar* 'fast/(to) fast', *dom/doar* 'gift/(to)give' (but *donativo* 'donation'), *ladrão/ladroagem* 'thief/thievery', *ordem/ordeiro* 'order/peaceful' (but *ordenar* '(to) order'), *tom/toar* 'tone/(to) sound' (but *tonal* 'id.'), *lua/lunar* 'moon/lunar', *um/uma* 'indefinite article-MASC/-FEM', *desdém/desdenhar* 'disdain/(to)disdain'. These forms are probably not to be synchronically related. The deletion of intervocalic *-n-* was a general phenomenon in Xth/XIth century but it was lost after the XIIth century (cf. Williams 1938; Teyssier 1980). Thus, these words can be assumed to have entered the language in different periods of time. See, nevertheless, d'Andrade (1977), d'Andrade and Kihm (1988) and Mateus and d'Andrade (2000: 7.2) for alternative views.

¹⁰² As said above, in Mateus (1975) and Mateus and d'Andrade (2000) the glide is inserted. We believe a simpler analysis is to allow for the underlying nasal segment to be specified as a non-syllabic nasal segment, which may surface as a nasal glide in word final position.

¹⁰³ Notice that fusion must precede nasal vowel rise (cf. Mateus 1975: 1.4), for otherwise a low central vowel would surface, as it happens across words (see also section 3.12). This, again, suggests that the realization of the nasal segment is a lexical phenomenon.

¹⁰⁴ The fact that there are some cases where the nasal segment surfaces as a consonant rather than a glide, as in (i), can be analyzed as involving an exceptional nasal consonant that is fully specified underlyingly.

- | | | |
|-----------------------|---------------|-----------------|
| (i) abdómen 'abdomen' | cólon 'colon' | íman 'magnet' |
| dólmen 'dolmen' | cláxon 'horn' | slogan 'slogan' |

Under this view, these cases do not constitute an exception to final nasal diphthongization.

¹⁰⁵ The same rule is also proposed in Andrade (1980: 57-58).

¹⁰⁶ We assume that this process applies within a phonological domain, since it does not seem to refer to any morphological information.

¹⁰⁷ Vowel harmony is a process specific for the verbal system that consists of the regressive long-distance assimilation of the height features of the theme vowel by the final vowel of the verb root. It only occurs when the theme vowel is deleted (e.g. Mateus 1975; d'Andrade 1994d; Mateus and d'Andrade 2000).

¹⁰⁸ Although Mateus and d'Andrade (2000) do not analyze /e/-centralization in detail, they assign it to the postlexical component.

¹⁰⁹ The syllabification of this glide contrasts with syllabification of the glide that is inserted to break a hiatus formed by oral vowels, as in *areia* 'send' (see section 3.6 below). In the latter case, according to Mateus and d'Andrade (2000: 57) the glide belongs to the rhyme of the syllable headed by /e/ (see also Gonçalves Viana 1883, for the syllabification of words like *maior* (mai.or) 'bigger'). Only in the latter case, therefore, the palatal glide provides the context for tautosyllabic /e/-centralization.

¹¹⁰ In the inflected forms inside brackets given in (48), the front vowel is stressless and there is no glide insertion. Instead, the front vowel followed by another vowel semivocalizes. This is also shown in (49) below.

¹¹¹ As reported in section 3.5, an alternative pronunciation of the forms in (48) admits a glide. However, this glide behaves differently in some respects: (i) it is totally optional, (ii) it does not trigger /e/-centralization, (iii) it belongs to the onset of the syllable that follows /e/, rather than to the preceding coda.

¹¹² We would like to thank an anonymous reviewer for bringing this point to my attention.

¹¹³ As noted in Andrade (1980: 4.2.1), this is true for non-round vowels. Round vowels show more variability, as illustrated in (i).

- (i) *emoldurar* [u]/[o] '(to) frame' (*mO*lde [ɔ] 'mould')
 soltar [o] '(to) set free' (*sO*lta [o] '(he) sets free')

In order to account for the general realization of unround vowels in syllables closed by /l/ as low, Andrade proposes that the lowering rule also applies in these cases.

¹¹⁴ We should point out that schwas are often deleted in EP (see in particular chapter 7 of this book).

¹¹⁵ Notice that this may be conceived as a feature specification process, rather than a feature changing one, since non-high non-central vowels in this position may be underlyingly underspecified for the low feature. We should also add that there is no neutralization in the case of central vowels, since non-high central vowels are underlyingly low.

¹¹⁶ Notice, nevertheless, that d'Andrade (1977) assumes that words ending superficially in a liquid have an underlying final /e/.

¹¹⁷ To be more accurate, we believe that there are two possible syllabifications for these sequences: *pVrs* or *prVs*, where V stands for a V-position that may be filled by [i]. In the first case, the sequence [rʃ] belongs to the rhyme, whereas in the

second, [r], like [p], belongs to the onset, and [ʃ] belongs to the rhyme (see also the remarks made in Morais Barbosa 1965: 5.16). What we believe is impossible is the separation of this sequence in two different syllables, as would be the case if an empty V-position existed between each of these consonants (*pV.rVs.). The latter syllabification seems to be allowed in Mateus and d'Andrade's proposal on syllable construction. External evidence for our claim comes from the difficulty speakers show in writing such words. While the alternation between the spellings *perspectiva* 'perspective' and *prespectiva* is frequent, the spelling *perespectiva* does not occur (unlike in other cases where a V-position may be posited, as in *segmento/seguemento* 'segment', cf. Vigário and Falé 1994).

¹¹⁸ /r/ behaves differently from other segments that appear in coda (i.e. /l/, the nasal segment, and /s/): /r/ undergoes strengthening processes in word initial position and in the initial position of syllables preceded by a consonant (cf. section 3.8); /r/ does not semivocalize in word final position (as the nasal segment, cf. section 3.4), or in word final position followed by a fricative (as the lateral and the nasal segment, cf. section 3.4); /r/ in word final stressed syllables may often delete (unlike lateral consonants) (cf. Leite de Vasconcellos 1901: 98). As there is an important generalization that there is only one consonant per rhyme in EP (e.g. Mateus and d'Andrade 2000), it follows that /r/ seems not to pattern like a 'full' consonant. This topic deserves further attention. See also Freitas (1995) for the problems involved in the acquisition of trills in EP.

¹¹⁹ Even if the second schwa is deleted, there is a clear perceptual effect concerning syllable structure: instead of the syllable division *li.dVrs*, if the vowel reduces speakers hear *li.drVs*. That is, in this case again the vowel does not behave as belonging to a syllable closed by /r/ in word final position.

¹²⁰ For the sake of simplicity, from now on we will refer to the uvular consonant in its trill realization ([ʀ]), rather than in its fricative realization ([ʁ]).

¹²¹ Notice that the "onset rule" does not apply to all cases where /r/ appears in onset position. In the second position of a consonant cluster, as well as in intervocalic position, /r/ is subject to the default process, rather than to the "onset rule".

¹²² The absence of sonorant palatal consonants preceded by /r/ deserves further investigation.

¹²³ Besides the impossibility of syllables closed by /r/, /l/ and /s/ in prefinal syllables of proparoxiton words, syllables with diphthongs and nasal vowels are also disallowed in post-tonic non-final position (see also Reighard and Almeida 1983, cited in Morales-Front and Holt 1997).

¹²⁴ Since we found no cases of morphological conditioning of this process, we assume it applies at the left-edge of the prosodic word.

¹²⁵ In a study on the acquisition of trills in EP, Freitas (1995) shows that the problems related to the production of these segments are not solved until the last

stages of phonological development. In addition, Freitas reports cases where word initial [R] is substituted for [r] (e.g. *rei* 'king', *rato* 'mouse'). Although we believe that child production data do not necessarily mirror the adult system, and also despite the fact that in other positions the trill may be substituted for other segments (namely, [ʎ], [l] and [j]), these facts may lend further support for the derived status of word initial [R] in EP.

¹²⁶ In the examples, we also show cases of English [r] in other positions of the word, where it is commonly identified with the EP alveolar tap.

¹²⁷ These facts further suggest that the nasal segment is basically linked to the coda, before it is associated to the nucleus of its syllable. Thus, it is probably not basically a non-attached floating autosegment, as proposed in d'Andrade and Kihm (1988).

¹²⁸ Glides in falling diphthongs are shown in d'Andrade and Viana (1994) to be part of the nucleus rather than of the coda position. An important argument in favor of this analysis is the fact that glides in falling diphthongs are nasalized along with the preceding vowel when followed by a nasal segment in the same syllable. The assumption that the glide is part of the nucleus and that the nasal segment spreads to this syllabic position, as proposed in d'Andrade and Kihm (1988), nicely accounts for these facts.

¹²⁹ This issue is investigated from a sociolinguistic point of view in Mascarenhas (1997) (and more recently in Rodrigues 2001). Mascarenhas shows that several sociolinguistic factors correlate with the specific realizations of word initial vowels (e.g. 'sex', and 'region'). However, no integrated phonological analysis is reached. We will describe the process following our intuitions, which match in general the realizations of Mascarenhas' speakers that belong to the same sociolinguistic group as ourselves. In Mateus (1975: 4.2) and Andrade (1980: 4.5), different descriptions of initial vowel realization can be found. Although not all the facts described in this section are given in Mateus and d'Andrade (2000: 3.3.2), our description generally agrees with the one found in this work.

¹³⁰ The description of initial vowels realization found in Gonçalves Viana (1883) points in this direction: front vowels are reported to always surface as high, while today they can also surface as mid; and initial back vowels in closed syllables are characterized as tending to be low, while this tends not to be the case in present day EP. In Leite de Vasconcellos (1901), in addition, round vowels are reported to tend to surface as high, regardless of dialect considerations, while they usually surface as mid or low in present day EP spoken in the region of Lisbon. See also Morais Barbosa (1988) for an overview of different descriptions of the realization of initial vowels since the 18th century.

¹³¹ Forms like [ɛ]dito-[e]/[i]dital 'edict/decreet' or [ɛ]timo-[e]/[i]timológico 'etym/etymological' seem to suggest that front low vowels may in some cases raise to [i]. However, in these forms a low stressed vowel does not necessarily imply that there is an *underlying* low vowel. In fact, as pointed out by Wetzels (1992)

non-high stressed vowels of proparoxiton words are subject to a lowering rule and thus these vowels typically surface as low. This rule appears to apply without exceptions to stressed vowels in word initial positions of proparoxiton words. Therefore, both *édito* and *étimo* may be assumed to have an underlying /e/. Under this analysis, the alternative realizations in stressless position are the expected pattern if raising in one degree optionally applies to /e/.

¹³² The stressed mid vowel in *ovo* and *olho* is historically the result of metaphony, which may be described as a process of long distance assimilation whereby a low vowel in stressed position becomes mid if the following (adjacent) word final syllable ends in *-o* (for a detailed analysis, see d'Andrade 1994c, and references therein). The fact that the mid vowel in *ovito* and *olhito* occurs in non-stressed position and is followed by a non-final syllable, and the fact that the low realization is impossible suggests that the vowel has been reanalyzed as underlyingly mid. Notice, furthermore, that metaphony is no longer a productive rule. We therefore assume that the mid vowel is present in the underlying representation of these forms. Alternatively, we have to assume that the process that explains the realization of these vowels follows metaphony.

¹³³ Given the lack of evidence that the initial position is defined morphologically, we assume that the relevant domain for the application of this rule is phonological (see also the discussion further below). We should further point out that the rule in (75) may, in fact, be seen as a trivial instance of Vowel Reduction. That is, unlike front and round vowels, central vowels pattern word initially in the same way as word internally.

¹³⁴ We are assuming the feature geometry proposed in Mateus (1997b) and Mateus and d'Andrade (2000), developed within Clements (1988) and Clements and Hume (1995) framework. For the basic tenets of autosegmental phonology see, for example, Goldsmith (1991). We should stress that we are also assuming that the processes under observation apply to vowels in the initial position of a prosodic domain. This issue is discussed below.

¹³⁵ In fact, the delinking of the [+low] feature may also affect the central vowel (and therefore no specification regarding the place feature is required in the formalization of the process). However, this vowel does not show the alternation between low and non-low realizations. Since they are obligatorily specified as [-low], in this case feature specification should apply in the lexical level, unlike in the case of non-central vowels.

¹³⁶ The postlexical feature specification is possibly conditioned by sociolinguistic factors (see Mascarenhas 1998; Rodrigues 2001).

¹³⁷ In some dialects, the items in (86b) are usually produced with two vowels instead of a falling diphthong (see, for example, Gonçalves Viana 1883; Sá Nogueira 1938). This fact also suggests that the glides in (86) are obtained by a process distinct from the one involved in the forms in (85).

¹³⁸ D'Andrade (1998) claims that in a sequence VV where the first one is stressed and the second is high, the second vowel always surfaces as a glide. He notes

furthermore that, in some dialects, the second vowel may surface as a vowel, but he claims that this is specific of the verbal system and is dependent on a tense morpheme boundary. However, the examples presented above involving a noun, and the forms involving a verb and a clitic (as in *vi-o* ‘(I) saw-him’), show that the second vowel may surface as a glide or as a vowel in prosodic word final position, regardless of morphological considerations (see also section 5.3.3.1).

¹³⁹ As mentioned in Câmara (1972), Gonçalves Viana (1883) presents (near) minimal pairs, such as *paulAda* ([aw]) ‘blow’ < *pAu* ([áw]) ‘stick’ and *paulAda* ([əu]) ‘region of swamps’ < *paUl* ([əú]) ‘swamp’, which are revealing in this regard.

¹⁴⁰ In fact, it is not impossible to conceive that the schwa surfaces as a glide, but in our dialect typically as a central but not a front one (a glide that exists, for instance, in Dutch—cf. L. Wetzels, personal communication). Nevertheless, in all the EP descriptions we are aware of, such a glide was never reported. We believe that the realization of the stressless final front vowel preceded by another vowel deserves further investigation.

¹⁴¹ Similar facts are also reported by d’Andrade and Viana (1993: note 3).

¹⁴² The rule is argued to apply within a *compound intonational phrase* (I^{max}). This domain is defined as a constituent that dominates two constituents of the same type (I) (see Frota 2000a for the arguments that support the existence of such a domain in EP, and section 1.4 of this book). We should notice that the upper domain that bounds this process is not crucial for our purposes here. The conditions for V1 semivocalization blocking, related to the presence of different levels of stress in the second vowel, are discussed in Frota (1996, 2000a) and Ellison and Viana (1996). This issue is not relevant for us either and thus will not develop it here.

¹⁴³ See d’Andrade and Viana (1993), Frota (1996, 2000a) and Ellison and Viana (1996) for further illustration.

¹⁴⁴ Notice that in cases such as *cuidadoso* ([uj]) ‘careful’, which has a falling diphthong, the second segment never occurs in a context where it surfaces as a vowel, and the glide is obligatory. In these cases, therefore, there is no reason to consider that the glide is obtained through V2 semivocalization.

¹⁴⁵ On verbs ending in an underlying /i/, see section 3.2.

¹⁴⁶ Although we have not investigated this issue in depth, our view of these facts is similar to the one presented for the rising phenomena that affect word initial stressless vowels. Specifically, we believe that non-high front stressless vowels (that is /e/ and /ɛ/) undergo a neutralization process that may be explained by the delinking of the features [low] and [high]. The specification of these lexically underspecified vowels as [+high] may be postlexical. As for centralization, the simplest analysis is to assume that it is also a postlexical phenomenon. This is so, because when these underlying vowels occur in stressless positions followed by other vowels, they usually surface as a non-back glide ([j]), as a consequence of V1 semivocalization (similarly to /i/). Thus, if they are lexically specified as non-back,

the only feature that has to be modified is the syllabic one. This issue is left for future research. In the formalization of the rule below, we mark these segments as [-high] in order for the rule not to apply to segments that are lexically specified as [+high] (i.e. /i/).

¹⁴⁷ To mention just two of the differences between Gonçalves Viana's (1883) data and most subsequent descriptions: Gonçalves Viana reports that a glide could surface if the following word starts with a stressless vowel, while this is not allowed in similar cases today; the presence/absence of the glide was dependent to some extent on V2 vowel quality, while today V2 quality does not seem to play any role in non-back non-high vowel deletion.

¹⁴⁸ The judgments we provide in this section come from two sources: our own intuition and the observation of the production of four speakers in a reading task. The description of this task is presented in chapter 7. The most important difference between our description and some of the more recent observations on non-back non-high vowel deletion is that for us (and our informants), there is, in general, no glide whether the second vowel is stressed or not (as in 98b). Although judgements similar to ours are provided by Ellison and Viana (1996), the realization of the glide when V2 is stressed is given as the general case in d'Andrade and Viana (1993).

¹⁴⁹ A different proposal is made in Ellison and Viana (1996) in a brief note on this issue. These authors suggest that when a glide surfaces it is the result of an insertion process. In Vigário (1998b: 4.1) we have extensively argued against this view, showing that the front vowel that may surface as a glide must be present underlyingly.

¹⁵⁰ We should add that when the non-high front vowel is preceded by a vowel, deletion usually does not apply. Rather, a very short schwa-like segment is articulated (see note 60).

¹⁵¹ Notice that the possibility of clitics to surface without a glide will be argued in chapter 7 to follow from a more general property of clitics, namely the fact that they undergo phonological reduction.

¹⁵² On the prosodic conditions for V1 semivocalization blocking, see Ellison and Viana (1996) and Frota (1997, 2000a). In these works, different analyses are also provided for the conditions on the blocking of round vowel deletion.

¹⁵³ The data in Frota (2000a), reproduced in (108), are based on actual productions of speakers in a reading task.

¹⁵⁴ Notice that in cases such as (108d), Frota (2000a) argues that the two I-phrases may be grouped together in a compound prosodic domain, and this is the domain within which the rule applies (on the notion of compound prosodic domain see also section 1.4 and note 77 below).

¹⁵⁵ The minimal clash configuration is defined as a sequence of word stresses which may include one intervening unstressed syllable (Frota 2000: 91-92).

¹⁵⁶ Vowel merger is described, among others, in Gonçalves Viana (1883), Sá Nogueira (1939), d'Andrade and Viana (1993), Ellison and Viana (1996), and

Frota (1997, 2000a). The upper domain that bounds this process is defined as the (compound) intonational phrase in Frota (2000a), where its sensitivity to stress-clash configurations is also assessed. The span nature of the process is proposed in Vigário (1999a).

¹⁵⁷ I^{\max} is defined as an intonational phrase that is dominated by a higher-level constituent. It may, however, dominate two constituents of the same level, in which case it forms a compound intonational phrase. This compound phrase is also reported to be the domain that bounds other rules. In addition, several facts show that the intonational phrases that are included in a compound phrase have also the properties of I-phrases: for instance, they also bound other phonological processes and set up the domain for tone assignment (see Frota 2000a for the complete argumentation and section 1.4 of this book for a review).

¹⁵⁸ Notice that, in the cases illustrated in (114), Sá Nogueira (1938), like ourselves, does not consider V2 semivocalization possible either (see section 3.10.1). We should add that, in any event, these particular issues are not crucial for our analysis. In chapters 5 and 6, we will use this process as a diagnostic for the prosodic structure assigned to a number of constructions, considering only the segmental contexts where there is agreement in the literature on the possibility of central vowel deletion.

¹⁵⁹ Some specific (very frequent) combinations admit syllable degemination with syllables headed by [ɐ] (cf. *Avenida da...* ‘avenue of...’).

¹⁶⁰ As it stands, we believe the conditions in (126) are capable of predicting the most clear cases of syllable degemination. Nevertheless, Mateus and d’Andrade (2000: 145) consider that in the sequence in (i), which violates the condition (iii), syllable degemination is possible.

(i) *disse sinceramente não* ‘(he) said sincerely no’

According to our intuitions, there is a strong tendency in these cases for the two consonants to surface as a long consonant, and thus what we have here is not syllable degemination but rather a sequence of two identical consonants that become adjacent as a consequence of schwa deletion. The same tendency is also found when a round vowel is involved, as illustrated in (ii).

(ii) *pouco capaz* ‘not-very capable’
posso saber ‘(I) may know’

We should further notice that Frota (2000a) also reports the possibility of syllable degemination in a case with a stressed [o] in the second syllable (in the sequence *campo pouco* [pupó]/[pó] ‘field not-very’). Despite the variations in the judgments of these case, what seems clear is that a non-high vowel is totally disallowed in the first syllable, whereas in the second syllable such a vowel is possibly tolerated.

¹⁶¹ Besides the conditions related to the phonological composition of the syllables involved in this process, Mateus and d’Andrade (2000) suggest that the presence of

phrasal stress blocks syllable degemination. However, they do not elaborate on this issue.

¹⁶² In this preliminary study, Frota and Vigário (2000) analyze the prominence transcriptions of 80 sentences of EP, which were perceived by each of the authors (see section 5.3.3.2 for further details).

¹⁶³ Although both d'Andrade and Viana (1989) and Frota and Vigário (2000) do not report any occurrence of the initial stress in the second syllable, this seems to be a possibility, although a much less common one, according to our informal observations of spontaneous speech. In d'Andrade and Laks (1992), d'Andrade and Viana (1989) and (1999), and Pereira (1999) the alternation of stresses occurring in the first and the second syllable of a word is also described for words with an odd number of pretonic syllables. According to these authors, this follows from the interaction of two distinct mechanisms of non-primary stress assignment. The stress found in the first syllable is obtained by the positionally determined non-primary stress, while the stress on the second syllable is determined by the alternating rhythmic stress that is computed from the primary stress leftwards. The fact that all the rhythmic stresses between the second syllable and the primary stress are not perceived, however, suggests that the stress in the second position of the word is possible in some cases because it still counts as word initial. See section 5.3.3.2 for further discussion.

¹⁶⁴ D'Andrade and Viana (1999: 91) claim that the initial position refers to the intonational unit ("unidade entoacional") (on this notion, see section 2.3.2). They further argue that the domain of the initial stress rule is the intonational phrase rather than the word. In section 5.3.3.2 we challenge this view and propose that the domain of the initial stress rule is the prosodic word.

¹⁶⁵ In the examples, we mark the syllables bearing emphatic stress with capital letters and underlining. Collected data on emphatic stress is presented in Appendix I.

¹⁶⁶ In these examples, boldface indicates this specific type of stress.

¹⁶⁷ Relying on our intuition and perception, we would characterize the initial emphatic stress as consisting of a local F0 rise on the initial syllable. In Vigário (1997a, 1998a) another type of stress is also reported to affect the negative adverb *não* 'not'. Since the emphatic stress falls on the initial position of words, and given that the negative adverb is monosyllabic, it is not clear whether the two phenomena are in fact the same. There is at least one difference that we can identify at this point between the two: in the case of the F0 rise in *não*, the F0 may remain high in the realization of subsequent pitch accents (cf. Vigário 1997a, 1998a), while we believe that in the case of initial prominence the F0 rise is necessarily a local phenomenon, affecting only the word initial syllable.

¹⁶⁸ It is possible that there is a correlation between initial stress and emphatic stress. That is, it can be the case that emphatic stress is, in fact, a stronger (emphasized) realization of initial stress. Nevertheless, it is possible to have emphatic stress in a syllable that is immediately followed by the main stress of a word, which we

believe is impossible for non-emphatic initial stress. This suggests these are two independent phenomena.

¹⁶⁹ The pitch accents (that is, the tonal events that are associated with stressed positions) that are well established for EP are L*H, HL*, and H*L (see, in particular, Frota 1994b, 2000a; Vigário 1998a). These phonological categories are related to the following phonetic properties: the vowel that is assigned an H is realized with a relatively high F0 value, whereas a vowel that is assigned an L is realized with a relatively low F0 value. The asterisk indicates that the preceding L or H is aligned with the stressed syllable. For the basic tenets of intonational theory this description is based on see, among others, Pierrehumbert and Beckman (1988) and Ladd (1996).

¹⁷⁰ PA stands for pitch accent. Examples of the contours that instanciate the possibilities in (132bi) and (132bii) are, respectively, the sentences (77) and (76) of Vigário (1998a: 228).

¹⁷¹ The precise status of this initial H tone as a pitch accent or an edge tone is still an open issue. It seems to signal I-initial position, but, instead of being associated to the left edge of the I-constituent, it aligns with the right edge of the leftmost constituent of I. Putting aside the question of what the best analysis for this tone is, the initial H and its alignment with the post-tonic syllable of the sentence initial constituent are usually described in the literature.

¹⁷² The prosodization of clitics like *as* in (133) is discussed in chapter 5.

¹⁷³ Earlier observations on the intonational properties of sentences with focused elements in EP can be found in Viana (1987: 4.3). Notice, nevertheless, that the notion of focus in Viana's work cannot be fully identified with the one presented in this section. For example, the contours (b-c) of the sentence "Quem deu um livro à Maria foi o Vasco" ('It was John who gave a book to Mary') in her Figure 2.18, are argued by the author to show three focused constituents in each case (cf. Viana 1987: 152), while, according to the characterization of focus made here, those contours could not be so analyzed.

¹⁷⁴ On the notions of focus see, in particular, Jackendoff (1972), Rochemont and Culicover (1990), and the review in Frota (2000a: 1.4).

¹⁷⁵ In the examples, the focused word is signaled with small caps.

¹⁷⁶ The phonetic details of the alignment of the H*L pitch accent, as opposed to the HL* pitch accent, are systematically investigated in Frota (2000a: 5.3.2) and (2000b).

¹⁷⁷ The domain of range compression after the focus accent is suggested in Vigário (1998a: 148) to be the intonational phrase.

¹⁷⁸ See note 88 on some basic notions of intonational phonology used here. We should add that the subscript 'i' identifies a boundary tone, that is a tonal event that is aligned with the segmental material at the edges of prosodic constituents.

¹⁷⁹ In the case of heterosyllabic /e/-centralization and glide insertion to break a hiatus, no direct evidence was found for the process to apply with reference to the prosodic domain. In these cases, we have followed the intuitive idea that, in the

unmarked cases, phonological rules apply within a domain defined by phonology, rather than by morphology. Recall also from sections 3.8 and 3.9 that /r/-strengthening and the particular realization of initial vowels are assumed to follow from strength properties of prosodic word initial position.

Chapter 4

On the affix/clitic status of stressless pronouns

1. Introduction

It has been widely observed that pronominal clitics in Romance languages display affix-like behavior (e.g. Klavans 1985; Zwicky 1987a; Spencer 1991; Halpern 1995; Kaiser 1999; and Crysman 1999, 2000a, 2000b). Among the properties pronominal clitics may share with lexical affixes are the following: (i) high selectivity with respect to the category of the host/base they attach to; (ii) special phonology triggered by the clitic-verb sequences; and in some cases, (iii) the ordering of clitics with respect to inflection. One way of accounting for these types of effects is to claim that pronominal clitics behave like lexical affixes because they are (inflectional) affixes. This position has been taken by Zwicky (1987a), Halpern (1995), and Kaiser (1999) for European Portuguese, by Monachesi (1996) for Italian, and by Miller and Sag (1997) for French.

Under the assumption that pronominal clitics are lexically combined with their hosts, what remains to be accounted for is the bulk of properties that clitics may share with *words*: for example, the ability of clitics to appear in different positions with respect to the verb, or the lack of lexical phonological interaction between verbs and clitics. In fact, the occurrence of such phenomena justifies the adoption of the opposite thesis, according to which pronominal clitics are postlexically inserted. The latter approach has been followed by Peperkamp (1997) for Italian, by Watson (1997) for French, and by Van der Leeuw (1997) and Vigário (1999a) for European Portuguese.¹⁸⁰ The issue is, however, not always systematically investigated, and in many cases no substantial evidence supports the specific view that is adopted.

For the purposes of this investigation, it is of crucial importance to determine whether EP pronominal clitics attach to the verb within the lexical component, or, on the contrary, are combined with their host postlexically. Specifically, if pronominal clitics are lexical affixes, we should expect them to behave phonologically like other inflectional affixes. In such a case, according to the facts presented in chapter 5, EP would show no cases of encliticization, and pronominal cliticization would simply become a trivial case of inflection. If, by contrast, pronominal clitics attach to their

hosts postlexically, we should expect them to behave phonologically like other postlexical clitics, such as prepositions and determiners. In such a case, any asymmetry between verbal enclitics and proclitics becomes an interesting fact for a theory of mapping between syntactic structures and prosodic structures, and for the discussion on the relevant information for prosodic word construction (see chapter 5, sections 4 and 6).

We will, therefore, review the facts of EP pronominal clitics in this chapter. We will show that the hypothesis that pronominal cliticization in this language is a postlexical operation is empirically superior to the hypothesis that pronominal clitics attach to their hosts in the lexical component. In addition, we propose an account of the apparent marks of lexicalization that is compatible with such a view.¹⁸¹

2. Potential evidence for the lexical attachment of pronominal clitics

It is often noted that the distinction between clitics and affixes is not always clear-cut (e.g. Zwicky and Pullum 1983, 1985; Spencer 1991; Anderson 1992; Halpern 1995, 1998). Zwicky and Pullum (1983: 503–504) list a number of properties that tend to distinguish affixes from clitics, and that we reproduce below:¹⁸²

- Clitics can exhibit a low degree of selection with respect to their hosts, while affixes show a high degree of selection with respect to their stems;
- Arbitrary gaps in the set of combinations are more characteristic of affixed words than of clitic groups;
- Morphophonological idiosyncrasies are more characteristic of affixed words than of clitic groups;
- Semantic idiosyncrasies are more characteristic of affixed words than of clitic groups;
- Clitics can attach to material already containing clitics, but affixes cannot.

Notice that most of these properties are presented in a non-categorical fashion.

Three of these properties have been invoked to argue in favor of the hypothesis that pronominal clitics in EP attach to their host within the lexical component. We present each of them in the next sections.

2.1. Specialized host

Like in other Romance languages, pronominal clitics in EP attach to hosts belonging to a specific morphosyntactic category – the verb. Since this sort of selectivity is also characteristic of lexical affixes (Zwicky and Pullum 1983), this property can be seen as a symptom of the affixal status of pronominal clitics, as suggested, for example, in Klavans (1985). Thus, for instance, pronominal clitics in EP, similarly to some derivational affixes (e.g. *-vel*, as in *adorável* ‘adorable’), and like all verbal inflectional affixes, may only attach to elements belonging to the category V, not N or A.¹⁸³

Conversely, this property distinguishes pronominal clitics from other phonological clitics in EP. For example, the definite article can be classified as a proclitic (see chapter 5), which can be hosted, within an NP domain, by words belonging to different morphosyntactic categories, as illustrated in (200).

- | | | |
|----------|---------------------------------------|---------|
| (200) a. | <i>a caneta</i> | ART–N |
| | ‘the pen’ | |
| b. | <i>a grande caneta</i> | ART–A |
| | ‘the big pen’ | |
| c. | <i>a então ministra da saúde</i> | ART–ADV |
| | the then minister of health | |
| | ‘the minister of health at that time’ | |

The high degree of selection of EP pronominal clitics with respect to their hosts is therefore a property these elements share with lexical affixes, and not with (other) clitics.

However, we will try to show in section 4.1 that this similarity is not necessarily evidence for the lexical attachment of pronominal clitics in EP (and in Romance languages in general), as this result may be obtained if the location of phrasal clitics is allowed to be established with reference to the head of the relevant syntactic constituent and not just to phrasal syntactic nodes.

2.2. Phonological idiosyncrasy

Phonological idiosyncrasy involving the sequence verb-clitic is another argument that has been used to support the lexical attachment hypothesis, since the locus of idiosyncrasy is usually accepted to be the lexicon (cf.

Zwicky 1987a; Halpern 1995; and chapter 1, section 3.1 of this book). In EP the sequence verb-clitic displays some special phonological behavior, as it will be shown in more detail in section 4.2.1. Two examples are provided in (201): (i) the accusative clitic, which otherwise has the form *o*, is realized as *lo* when it is preceded by a verb ending in a consonant (that is subsequently deleted), and it has the form *no* when it is preceded by a nasal diphthong (see 201a); (ii) the final consonant of the verbal host is deleted when followed by *nos* (see 201b).

- | | | | |
|----------|--------------------|-----|---------------------|
| (201) a. | <i>comes</i> | but | <i>come-lo</i> |
| | eat-PRES2Sg | | eat-PRES2Sg-3SgACC |
| | <i>como-o</i> | | <i>comem-no</i> |
| | eat-PRES1Sg-3SgACC | | eat PRES3Pl-3SgACC |
| b. | <i>damos</i> | but | <i>damo-nos</i> |
| | give-PRES1Pl | | give-PRES1Pl-1PlDAT |

Although this sort of data has sometimes been (implicitly) considered in the literature on EP to follow from pure phonological processes, such an analysis is in fact not tenable on synchronic grounds.¹⁸⁴ Indeed, there are no general phonological rules in EP that could derive these forms.

Under the hypothesis that pronominal clitics are combined with their hosts postlexically, the phenomena mentioned above needs to be accounted for (see also Kaiser 1999; Crysmann 2000a, 2000b). In section 4.2.2 we propose an analysis, along the lines of Hayes' (1990) precompiled phrasal allomorphy, that allows for an account of phonological idiosyncrasy compatible with the postlexical insertion of pronominal clitics.

2.3. "Inflection" after cliticization

The third argument that has been put forward in favor of the affixal nature of pronominal clitics in EP is the possibility of pronominal clitics to precede inflection. Many researchers have argued that this happens in the construction traditionally called *mesoclis* (e.g. Zwicky 1987a; Halpern 1995; Kaiser 1999; Crysmann 2000a, 2000b): as exemplified in (202), when a clitic follows a verb inflected for the future or conditional, the clitic is inserted before what is taken to be a person/number affix (Zwicky 1987a: 143), instead of appearing at the end of the inflected verb.

- (202) a. *perceberia* understand-COND3Sg
 falaremos speak-FUT1Pl
 b. *perceber-te-ia* (**perceberia-te*) understand-COND3Sg-2SgDAT
 falar-lhe-emos (**falaremos-lhe*) speak-FUT1Pl-3SgDAT

According to both Zwicky and Halpern, it is the lexical status of pronominal clitics in EP that explains why these elements may appear before inflection: since clitics are inflectional affixes it is predicted that they should be able to appear “inside of (other) inflections” (Halpern 1995: 186).

As we will see below, this argument crucially depends on the analysis of mesoclis. In opposition to the authors mentioned above, we argue in section 4.3 that what follows the clitic is not a lexical inflectional affix, and that mesoclis must be the result of a syntactic operation. Thus, we will sustain that mesoclis does not constitute a valid argument for the inflectional status of pronominal clitics in EP.

3. Evidence for the postlexical insertion of pronominal clitics

Despite the marks of lexicalization reported in the preceding section, there is also substantive evidence that pronominal clitics in EP must be independent of their hosts at the lexical level. The evidence includes not only the remaining properties of Zwicky and Pullum’s (1983) list, referred to above, but also additional facts that we believe can be used as further diagnostics for the clitic/affix distinction. We present the relevant data in the following paragraphs.

3.1. Distributional facts

EP pronominal clitics are manipulated by syntactic (or at least non-lexical) operations, as they may occur either pre- or postverbally, a pattern never found with lexical affixes, which either attach to the right or to the left of their base.¹⁸⁵

Moreover, the distribution of clitics depends on phrasal information: they are preverbal when preceded within a given domain by certain adverbs and quantifiers, Wh-operators, complementizers and negative words (see 203).¹⁸⁶

132 *On the affix/clitic status of stressless pronouns*

- | | | |
|-------|---|---|
| (203) | a. <i>dou-te</i>
give-PRES1Sg-2SgDAT
'I give you' | a'. <i>não te dou</i>
not 2SgDAT give-PRES1Sg
'I don't give you' |
| | b. <i>eles ouviram-te</i>
they hear-PAST3Pl-2SgDAT
'they heard you' | b'. <i>todos eles te ouviram</i>
all of-them 2SgDAT hear-PAST3Pl
'they all heard you' |

Since the syntagmatic information relevant for the distribution of pronominal clitics is not available in the lexicon, the ordering of the sequence V-CL must be obtained postlexically (the same point is also made in Spencer 1991 and Crysmann 2000a, 2000b).

Further, under the view that pronominal clitics are inflectional affixes, the very possibility of proclisis in an inflectional system such as EP's is unexpected, since in this language inflectional morphology is exclusively suffixal.¹⁸⁷

The possibility of interpolating elements between a proclitic and the verb, as illustrated in (204), also indicates that the sequence CL-V cannot be created in the lexicon and inserted as a unit at the moment of lexical instantiation.¹⁸⁸

- (204) a. *Gostaria que o **não** fizesses.*
 (I) would like that 3SgACC not (you) do
 'I would like you not to do it.'
- b. *Apenas te **não** digo uma coisa.*
 only 2 SgDAT not (I) say one thing
 'I'm only omitting one thing to you.'

Another argument in favor of the non-affixal status of pronominal clitics comes from the possibility of clitics to have scope over a conjunction, as noted in Matos (1997) and Rouveret (1999) (see 205). This type of behavior is not found with lexical affixes, inflection included.¹⁸⁹

- (205) *Todos **o** aplaudiram e festejaram.*
 all 3SgACC applauded and celebrated
 'They all applauded and celebrated it.'

Matos observes in addition that, contrary to what is commonly thought (cf. among others, Cunha and Cintra 1984; Rouveret 1999), this behavior is

found not only with proclitics but also with enclitics, as shown in (206) (from Matos 1997: 707).

- (206) a. *Eles tinham-nas ouvido às avós e contado aos filhos.*
 they had 3SgACC-FEM heard to-the grand-mothers and told
 to-the sons
 ‘They had heard them from their grand-mothers and told
 them to their sons.’
 b. *Ele estava-lhe sempre a telefonar e a comprar livros.*
 he was 3SgDAT always phoning and buying books
 ‘He was always phoning her and buying her books.’

As pointed out in Spencer (1991: 9.1.5), the fact that the clitic is not doubled by a full NP also adds to the non-inflectional nature of pronominal clitics in EP. Indeed, in languages with agreement markers like Chukchee (a language spoken in Eastern Siberia), the agreement affix must occur with the overt NP it agrees with. Thus, EP pronominal clitics do not behave at least like these agreement markers.

By the same token, since clitic pronouns may not co-occur with full NP’s in EP, clitic pronouns are unlike inflectional affixes in that they are not obligatory – as, for example, inflectional subject agreement markers are.

Finally, clitic climbing is another distributional argument that is often argued to show that pronominals do not behave syntactically as inflection. In fact, in some EP constructions, clitics that are selected by a lower verb can also attach to a higher verb, as illustrated in (207).

- (207) a. *queria dar-te um livro*
 b. *queria-te dar um livro*
 ‘I would like to give you a book’

However, as pointed out in Halpern (1998: 106), this clitic positioning may be a consequence of V-V restructuring. Restructuring has indeed been advocated for EP in Gonçalves (1994, 1999). Under such a view, this is not an instance of a (long) optional clitic movement, but rather a consequence of the presence/absence of restructuring. Thus, clitic climbing is possibly reduced to a trivial case of “local cliticization” (see, for example, Manzini 1998 for a review of this issue).

As for all of the other distributional facts reported in this subsection, it seems to us that they constitute sound evidence for the non-affixal status of pronominal clitics in EP.

3.2. *Phonological facts*

A number of phonological facts also call for a non-inflectional analysis of pronominal clitics. Actually, all the processes classified in chapter 3 as applying within the lexical component are not affected (triggered or blocked) by phonological information that results from the attachment of a clitic to the verb. This is expected under the view that pronominal clitics are not combined with their host in the lexical component.

First, clitics do not affect the location of verbal stress, as shown in (208), even if the addition of clitics causes the stress to fall on the fifth syllable (from right to left) of the verb-clitic sequence.

- (208) *diz[í]amos*
 tell-IMP1Pl
diz[í]amo-lo
 tell-IMP1Pl-3SgACC
diz[í]amo-no-lo
 tell-IMP1Pl-1PlDAT-3SgACC

Since otherwise stress may never fall further leftward than the antepenultimate syllable, we should conclude that pronominal clitics are not included in the EP stress window (see also Van der Hulst 1996 for the claim that the three syllable stress window is universal). A natural way of explaining this fact is to assume that pronominal clitics are not present in the string when word stress applies.

Moreover, the insertion of clitics does not block the process of word final nasal diphthongization, as example (209) illustrates: in the EP variety under observation we only find nasal diphthongization in word final position and never inside a word regardless of its internal structure (see chapter 3, section 5).¹⁹⁰

- (209) /e[+nas]/ word internally: *entediar* [ẽ]/*[ẽĩ] ‘to tire’
batente [ẽ]/*[ẽĩ] ‘door-knocker’
 /e[+nas]/ word finally: *batem* *[ẽ]/[ẽĩ] ‘(they) hit’
 /e[+nas]/ before a word: *batem todos* *[ẽ]/[ẽĩ] ‘(they) hit all’
 /e[+nas]/ before a clitic: *batem-te* *[ẽ]/[ẽĩ] ‘(they) hit you’

The existence of nasal diphthongization before clitics thus indicates that the clitic is not incorporated into the verb when this process applies.¹⁹¹

Further, clitics do not trigger heterosyllabic /e/-centralization, as illustrated in (210). This process centralizes a mid non-back vowel when followed by a heterosyllabic palatal segment (see chapter 3, section 6).

- | | | | | | |
|----------|----------------|----------|----|--------------------|----------|
| (210) a. | <i>abelha</i> | [ɐ]/*[ɛ] | b. | <i>dê-lha</i> | *[ɐ]/[ɛ] |
| | ‘bee’ | | | give-it FEM-to him | |
| | <i>espelho</i> | [ɐ]/*[ɛ] | | <i>lê-lho</i> | *[ɐ]/[ɛ] |
| | ‘mirror’ | | | read-it-to him | |
| | <i>telha</i> | [ɐ]/*[ɛ] | | <i>vê-lhos</i> | *[ɐ]/[ɛ] |
| | ‘tile’ | | | see-them-to him | |

The non-application of /e/-centralization in otherwise similar phonological conditions also indicates that clitics are not attached to their host when the rule applies.

Finally, clitics do not trigger glide insertion to break a hiatus. This process, illustrated in (211), consists of the insertion of a non-back glide between two vowels when the first one is a stressed /e/ (see chapter 3, section 7). The examples in (211a–b) show that it is word bounded, since it does not apply if the first vowel is in word final position. The absence of glide insertion when the following vowel belongs to a pronominal clitic, as in (211c), is again an indication of the non-lexical attachment of pronominal clitics.

- | | | |
|----------|------------------|-------------------|
| (211) a. | <i>passe[j]o</i> | <i>rece[j]o</i> |
| | ‘walk’ | ‘(I) fear’ |
| | <i>are[j]a</i> | <i>rece[j]a</i> |
| | ‘sand’ | ‘(he) fears’ |
| | <i>recre[j]o</i> | <i>rece[j]e</i> |
| | ‘playground’ | ‘(he) fears-SUBJ’ |
| b. | <i>vê o João</i> | <i>vê animais</i> |
| | ‘see John’ | ‘see animals’ |
| c. | <i>vê-o</i> | <i>vê-a</i> |
| | ‘see it’ | ‘see it(-FEM)’ |
| | <i>lê-o</i> | <i>lê-a</i> |
| | ‘read it’ | ‘read it(-FEM)’ |
| | <i>dê-o</i> | <i>dê-a</i> |
| | ‘give it’ | ‘give it(-FEM)’ |

Certain clitics have been shown to interact with lexical phonological rules in languages such as Polish (cf. Booij and Rubach 1987; Spencer

1991: 9.1.4), and this has been taken to follow from the lexical attachment of clitics to the verb (cf. Halpern 1995). An illustrative example is provided by some Polish auxiliaries and tense markers. Like EP pronominal clitics, these elements may appear in different sentence positions, as documented in (212), taken from Booij and Rubach (1987).

- (212) a. *Ja to robił-e-m*
 b. *Ja to-m robił*
 c. *Ja-m to robił*
 ‘I did this’

In these cases, however, clitics interact with lexical processes, since they affect the location of word stress, they block the application of a lexical rule that raises vowels in word final position, and they undergo a lowering rule that also has the properties of a lexical phenomenon. As we have seen above, EP does not present these types of effects. The lack of such behavior in this language not only remains unaccounted for if the same affixal analysis of clitics is extended to European Portuguese, but also makes wrong predictions about the application of lexical phonological processes.

Other lexical rules that clearly distinguish inflection from cliticization in EP are those that refer to morphological information (for example, those applying only in verbal or inflectional environments). One of these rules consists of the centralization to schwa of the theme vowel of verbs belonging to the 3rd conjugation (/i/) in stressless final position, as illustrated in (213a) (see chapter 3, section 3). This process of centralization is characteristic of verb final position, and thus it does not occur before inflectional affixes, as shown in (213b). However, it does operate when the vowel is followed by a pronominal clitic, as in (213c).

- (213) a. *part[i]* ‘(he) breaks’
 b. *part[i]remos* ‘(we) will break’
 c. *part[i]-me* ‘(he) breaks me’

This shows that when the rule applies the clitic is not present in the string, and, even more importantly, that the pronominal clitic is not being treated as an inflectional affix.

There is another process that applies in inflectional environments and that shows the non-inflectional status of pronominal clitics – the deletion of the theme vowel when followed by an inflectional affix starting with a

vowel (see chapter 3, section 4). Examples of vowel deletion can be found with a Person/Number affix, as in (214a), or with a Tense/Mood/Aspect marker, as in (214b).

- | | |
|---|--|
| (214) a. <i>como</i> (<com e+o)
eat-PRES1Sg
<i>falo</i> (<fal a+o)
speak-PRES1Sg | b. <i>coma</i> (<com e+a)
eat-SUBJ1Sg
<i>falemos</i> (<fal a+e+mus)
speak-SUBJ1Pl |
|---|--|

The fact that vowel deletion never applies when the vowel is followed by a pronominal clitic, as illustrated in (215), also adds to the bulk of evidence that pronominal clitics are not inflectional affixes.

- (215) *come-o* (**como*)
eat-PRES3Sg-3SgACC
fala-o (**falo*)
speak-PRES3Sg-3SgACC
come-a (**coma*)
eat-PRES3Sg-3SgACCFEM

To conclude, in the same way lexical rules that operate in inflectional environments are not affected by the presence of clitics, the phonological idiosyncrasy found in sequences involving clitics is also not found in inflectional environments (see section 4.2).

3.3. Other facts

Some additional facts follow naturally under the hypothesis pursued here that pronominal clitics in EP are not inflectional affixes.

Pronominal clitics are peripheral with respect to inflectional suffixes, that is, they appear after all inflectional suffixes (see 216).

- | | |
|---|---|
| (216) <i>fala-me</i>
speak-PRES3Sg-1SgDAT
<i>falava-me</i>
speak-IMP3Sg-1SgDAT
<i>falávamos-te</i>
speak-IMP1Pl-2SgDAT | (root+TV+clitic)

(root+TV+TMA+clitic)

(root+TV+TMA+PN+clitic) |
|---|---|

138 *On the affix/clitic status of stressless pronouns*

This is expected under the hypothesis that clitics attach postlexically to fully inflected words.¹⁹²

In addition, while portmanteau forms in EP may involve inflectional affixes, on the one hand, and clitics, on the other hand, there are no portmanteau forms involving both inflectional affixes and clitics.¹⁹³ Examples of the former type are the TMA and PN markers of the perfect indicative, as illustrated in (217a) (e.g. Villalva 1994), and examples of the latter type are the combinations of accusative and dative clitic, as illustrated in (217b) (see also section 4.2, below).

- | | |
|---|--|
| (217) a. <i>falei</i>
speak-PERF1Sg
<i>falaste</i>
speak-PERF2Sg | b. <i>dou-lho</i>
give-PRES1Sg-3SgDAT-3SgACC
<i>dá-ma</i>
give-PRES3Sg-1SgDAT-3SgACCFEM |
|---|--|

Furthermore, cliticization is never restricted to hosts with specific phonological characteristics. As shown in Carstairs-McCarty (1998), this sort of phonological constraints can be found both in derivation and in inflection. A classical case of affix sensitivity to the phonological properties of its base, is provided by the English derivational suffix *-er*: *-er* attaches to monosyllabic and some disyllabic bases, not to trisyllabic forms, as illustrated in (218), from Sproat (1988).¹⁹⁴

- | | |
|---|--|
| (218) a. <i>blaker</i>
<i>softer</i>
<i>happier</i> | b. <i>*eloquenter</i>
<i>*irascibler</i>
<i>*importanter</i> |
|---|--|

In EP, Villalva (1994: chap.5, note 33) reports the existence of two suffixes with the same function that attach to bases with different number of syllables. According to this author, the nominal suffix *-eza* tends to select monosyllabic or disyllabic bases, while the suffix *-ez* tends to select bases with more than three syllables. Each case is illustrated in Villalva with the examples given in (219a) and (219b), respectively.¹⁹⁵

- | | |
|--|---|
| (219) a. <i>crueza</i> (root: <i>cru-</i>)
'crudity'
<i>franqueza</i> (root: <i>franc-</i>)
'honesty' | b. <i>estupidez</i> (root: <i>estupid-</i>)
'stupidity' |
|--|---|

As we said earlier, no such cases are found with EP pronominal clitics.

Finally, there are no arbitrary gaps or idiosyncratic semantics in sequences involving verbs and clitics, contrary to what often arises with affixes (cf. Zwicky and Pullum 1983): EP pronominal clitics attach to any element with the label V without exceptions, and, as far as we know, cases of idiosyncratic semantics are not reported in the literature.¹⁹⁶

To sum up, the bulk of evidence presented in this section clearly justifies the rejection of the hypothesis that pronominal clitics are lexical affixes in EP. Our task is, therefore, to develop an account of the marks of lexicalization noted in the preceding section, which does not force an analysis of pronominal clitics as lexically attached elements.

4. On the marks of lexicalization

4.1. Selectivity with respect to the host

As we have seen, pronominal clitics in EP, like in other Romance languages, attach to verbs.

Based on a survey of an important number of languages, Klavans (1985) proposes a theory of clitic positioning and attachment, whereby cliticization is argued to always involve *phrasal affixation*. Clitics are thus seen to attach only to phrasal nodes of the XP level, or of the X' level: the former case can be exemplified by Ngiyambaa (an Australian language), where clitics attach to the S node; while the latter case can be illustrated by Navajo (a North American Indian language), where clitics attach to N' positions. This view leads to the conclusion that the case of Romance clitics, which attach at the level of X⁰, is exceptional. Observing that pronominal clitics in these languages have insertion requirements that resemble inflectional clitics, Klavans suggests that these clitics are becoming affixes. From what we have seen, however, this analysis cannot be adopted for EP, since there are many facts pointing to the phrasal nature of pronominal cliticization in this language (cf. section 3.1).

Although Klavans' theory of clitic attachment excludes cliticization at the level of X⁰, other proposals have been made that allow for a phrasal treatment of pronominal cliticization in the Romance languages. Such view is explored, particularly, in the work of Anderson.

Anderson (1992) proposes that affixation and cliticization are the result of the same kind of morphological operation, the former applying to words and the latter applying to phrases. Moreover, clitic placement rules may refer to the *head* of a syntactic constituent, namely the head of the sentence

– V. Under this approach, the selection of a host belonging to a specific morphosyntactic class is the result of the specification of a parameter that rules the distribution of phrasal affixes. Thus, this property does not necessarily imply that the relevant clitics are becoming lexical affixes.

There are other possible approaches to cliticization that do not imply a lexical analysis of pronominal clitics in Romance languages. For example, in a significant number of studies where pronominal cliticization in Romance languages is considered to follow from a syntactic operation, pronominal cliticization is treated as a phenomenon distinct from inflection.¹⁹⁷ One such study is Duarte and Matos (2000), where it is proposed that Romance pronominal clitics are generated as arguments of the verb, and not under a functional head, like inflection.

For Duarte and Matos (2000), who develop their analysis within the Minimalist Program (Chomsky 1995), Romance pronominal clitics have a strong formal feature – “V-host” – that has to be checked against a V-head, forcing them to move in Overt Syntax. Thus, under this view, it is the strength of the formal feature V-host (and also Case) that explains not only the different syntactic distribution of clitics with respect to full phrases, but also their attachment to a head position.

Many of the aspects related to the distribution of pronominal clitics in EP, presented above in section 3.1, suggest that pronominal cliticization is probably to be analyzed as a syntactic phenomenon, rather than a morphological operation.¹⁹⁸ However, this issue is clearly out of the scope of this chapter, and it is still a matter of debate. We will therefore not take a position on this topic here.

Regardless of what the best account of cliticization in EP is, it seems clear that there are several possible ways of understanding the attachment of clitics to a verb form that do not imply their lexical treatment. Indeed, from what we have seen so far, the relevance of host selectivity as an argument in favor of the lexical treatment of pronominal cliticization appears to be dependent on theoretical assumptions, such as those of Klavans (1985).

We, therefore, conclude that the attachment of pronominal clitics to a specified class of words is a property clitics share superficially with lexical affixes, but it is far from representing clear evidence for their lexical treatment.

4.2. Phonological idiosyncrasy

4.2.1. On the nature of the formal alternations affecting the verb and the clitic

As we have seen in section 2.2, the combination of verbs and clitics displays idiosyncratic phonological behavior. In the following paragraphs we describe the relevant facts in more detail.

Let us first consider the alternations in the form of the accusative pronoun, which may be *o*, *lo*, and *no*.¹⁹⁹ As shown in (220), the clitic appears as *lo* when the preceding element (the verb or another clitic) ends with a consonant; this consonant in turn is deleted when the clitic is added.

- | | | |
|----------|---------------------|----------------------------|
| (220) a. | <i>comes</i> | <i>come-lo</i> |
| | eat-PRES2Sg | eat-PRES2Sg-3SgACC |
| b. | <i>dar</i> | <i>dá-lo</i> |
| | give-INF | give-INF3SgACC |
| c. | <i>damos</i> | <i>damo-lo</i> |
| | give-PRES1Pl | give-PRES1Pl-3SgACC |
| d. | <i>dás-nos</i> | <i>dás-no-lo</i> |
| | give-PRES2Sg-1PlDAT | give-PRES2Sg-1PlDAT-3SgACC |

The form *no* appears when the preceding verb ends with a nasal diphthong (see 221), which, at least for some speakers, including ourselves, must also correspond to the 3rd P(erson)Pl(ural) suffix, as shown by the contrast in (222).²⁰⁰

- | | |
|----------|--------------------|
| (221) a. | <i>comem-no</i> |
| | eat-PRES3Pl-3SgACC |
| b. | <i>comam-no</i> |
| | eat-SUBJ3Pl-3SgACC |

- | | |
|----------|--------------------|
| (222) a. | <i>põem-no</i> |
| | put-PRES3Pl-3SgACC |
| b. | <i>põe-o</i> |
| | put-PRES3Sg-3SgACC |

The form *o* appears elsewhere (cf. 223).

142 *On the affix/clitic status of stressless pronouns*

- (223) a. *como-o*
eat-PRES1Sg-3SgACC
b. *dei-o*
give-PAST1Sg-3SgACC

These alternations only arise when the clitic is preceded by a verb or another clitic, but not by other elements, as shown in (224).

- (224) a. *apenas o dou* **apenas lo dou*
only it (I) give
'I only give it'
b. *não o dou* **não no dou*
not it (I) give
'I don't give it'

The demonstration that these forms cannot be obtained through the application of phonological processes with some productivity in Modern EP is rather straightforward. This task was undertaken in Vigário (1999b: 584–586), where it is shown that such forms must be considered instances of allomorphic variation. Since there seems to be little place for controversy on this issue, we adopt this view here without further discussion.

Let us now consider consonant deletion before certain clitic forms.

Consonant loss only occurs regularly in verb final position before accusative clitics, not before clitics marked for a different case (see 225a vs. b).

- | | |
|---|---|
| <p>(225) a. <i>dá-lo</i> (<dar-lo)
give-INF-3SgACC
<i>damo-lo</i> (<damos-lo)
give-PRES1Pl-3SgACC
<i>fá-lo</i> (<faz-lo)
do-PRES3Sg-3SgACC</p> | <p>b. <i>dar-nos</i>
give-INF-1PlDAT
<i>damos-lhe</i>
give-PRES1Pl-3SgDAT
<i>faz-te</i>
do-PRES3Sg-2SgDAT</p> |
|---|---|

It affects, in addition, specific combinations of verb and clitic forms: the verb final consonant is deleted (i) if it belongs to the 1st PPl marker (but the consonant is not deleted if the verb is in the 1st, 2nd, or 3rd PSg forms – see 226a vs. 226b), and (ii) if the following pronoun is the 1st or 2nd PPl dative clitic (but not if it is the 2nd PSg, 3rd PSg or 3rd PPl dative forms – see 226a vs. 226c).

- (226) a. *damo-nos* (<damos-nos) *damo-vos* (<damos-vos)
 give-PRES1Pl-1PlDAT give-PRES1Pl-2PlDAT
 b. *pus-nos* *faz-nos*
 put-PRES1Sg-1PlDAT do-PRES3Sg-1PlDAT
 dás-nos
 give-PRES2Sg-1PlDAT
 c. *damos-te* *damos-lhes*
 give-PRES1Pl-2SgDAT give-PRES1Pl-3PlDAT
 damos-lhe
 give-PRES1Pl-3SgDAT

It also affects specific combinations of clitics, namely *nos* and *vos* (but not *lhes*) followed by the accusative forms *lo(s)/la(s)* – see (227a) vs. (227b).

- (227) a. *deste-no-lo* b. *dou-lhos* (**dou-lhe(s)-lo*)
 give-PAST2Sg-1PlDAT-3SgACC give-PRES1Sg-3PlDAT-3SgACC
 dou-vo-lo
 give-PRES1Sg-2PlDAT-3SgACC

Again, it is clear that this behavior is not found elsewhere in the phonology of EP. As the examples in (228) show, these sequences of consonants are well-formed in EP, and are not subject to this specific process of consonant deletion.²⁰¹

- (228) a. *orla* b. *âmbar lindo*
 ‘border’ ‘pretty amber’
 eslavo *belas luvas*
 ‘Slav’ ‘beautiful gloves’
 asno *pires novo*
 ‘ass’ ‘new plate’
 desvendar *pires verde*
 ‘unravel’ ‘green plate’

Therefore, consonant deletion must be considered an idiosyncratic process (see further below).

4.2.2. Towards an account of the phonological idiosyncrasy in verb-clitic sequences

As we have seen, phonological idiosyncrasy affecting strings of morphemes is one of the properties listed in Zwicky and Pullum (1983) that tends to distinguish clitics from affixes in that it is more characteristic of combinations involving affixes than of combinations involving clitics. Further, Zwicky (1987a) and Halpern (1995) assign it a major role in their claim that EP pronominal clitics are inflectional affixes.

There are, however, various examples in the literature of phonological processes that occur between words and have lexical properties, i.e., they are restricted to apply only with specific items or classes of items. Numerous examples of such processes in different languages can be found, for example, in Hayes (1990), Kaisse (1990), Odden (1990), and Nespor (1990). Although some of the cases reported do involve the combination of a clitic with its host, many cases can also be found that do not involve such type of combination. For example, Hayes (1990) reports the case of Hausa (described in Kraft and Kirk-Green 1973), where a shortening process applies to final long vowels of verbs when followed by an NP direct object. Under the view that this type of behavior is an indication that the elements involved are combined in the lexicon, every instance of a verb-direct object NP sequence in Hausa would have to be analyzed as a lexical combination. This position is hard to maintain. Hayes provides other examples of rules that cannot be classified as purely phonological and that do not involve clitics: in Ewe, a particular tonal rule only applies if its phonological description is met and the target tone is followed by a noun root (cf. Clements 1978); in colloquial French, liaison occurs between adjectives and quantifiers, but not between an adverb and a verb or adjective (cf. Selkirk 1972); in Mende, there is a rule that lenites different word initial consonants, but only in certain syntactic environments (cf. Cowper and Rice 1987). Kaisse (1990) refers to the case of Modern Greek, where there is a rule that deletes a word final vowel when the following word starts with a vowel. This rule only applies in certain syntactic conditions: it has to mention category labels and syntactic relationships. Odden (1990) reports the case of Kimatumbi, where a number of processes also refer to syntactic information. Nespor (1990) adds the case of *Troncamento* in Italian, a process of word final vowel deletion which is shown to apply productively only with verbs.

These cases show that phonological idiosyncrasy associated with the combination of specific items or classes of items is not restricted to se-

quences of verbs and clitics. More importantly, they can be seen to demonstrate that phonological idiosyncrasy associated with certain combinations of words does not necessarily imply that those words are combined in the lexicon.

In order to account precisely for this kind of phenomena, Hayes (1990) develops a theory of precompiled phrasal allomorphy. This theory provides a framework that will enable us to analyze as a lexical phenomenon the phonological idiosyncrasy found in the combination verb-clitic, without having to assume that verbs and clitics are combined in the lexicon. This seems the best approach since, as we have shown, on the one hand, the alternations in the form of the verb and in the form of the clitic do not follow from pure phonological rules of EP, and, on the other hand, there is compelling evidence for the non-lexical combination of verbs and clitics in this language.²⁰²

As we have seen in more detail in chapter 1, section 3.1, precompilation may consist of lexical listing of allomorphs, specified with the environment for their phonological instantiation, or it may consist of lexical rules that derive allomorphs and that refer to instantiation frames in their structural description. The instantiation frames, in turn, define the syntactic context for the insertion of the appropriate allomorph.

Adopting this framework, we propose to account for the alternations in the form of EP accusative pronominal clitics as instances of lexical listing, since these alternations are specific to these elements. In accordance with the distribution of each form described in section 4.2.1, the lexical entry of the accusative clitic pronoun should therefore include the information shown in (229).²⁰³

$$(229) \left(\begin{array}{l} \dots \\ no / [\dots]_{Vb} \text{ — } \\ \quad \quad [3^{rd}P] \\ lo / [\dots [+cons]]_{Vb} \text{ — } \\ o / (\text{elsewhere}) \\ \dots \end{array} \right)$$

Thus, the allomorph *no* is inserted whenever the preceding verb is inflected for the 3rd PPI, while *lo* is selected when the verb ends with a consonant. In the remaining contexts, the inserted form is *o*.²⁰⁴

As for the consonant loss, both in verb final position and in 1st PPI morpheme final position, we propose that it follows from a lexical rule,

formalized in (230), that refers to the instantiation frames formalized as in (231).²⁰⁵

(230) $C \rightarrow \emptyset$ / [... _ ...]_{[Frame 1]; [Frame 2]}

(231) Frame 1: [... _]_{Vb} [...]_{CL[ACC]}
 Frame 2: [... _]_{Vb[1PL]} [...]_{CL['nos'/'vos']}

Therefore, the verb final consonant is deleted in the verb form that is inserted in the syntactic context defined by the frames in (231).

Finally, in the case of the different forms involving sequences of clitics we propose that they do not follow from the application of specific rules. Rather, we believe that a more adequate analysis is to treat the EP sequences of clitics as lexically formed clusters. This approach has been followed for different languages, such as Bulgarian and Serbo-Croatian (Halpern 1995), Sekani (an Athabaskan language – Halpern 1996), Italian (Monachesi 1996), French (Watson 1997), among others, on the basis of some type of specific behavior of clitic clusters. The same kind of specificity involving clitic clusters is found in EP:

- (i) clitics behave differently in similar phonological environments: the combination of *nos*+(*l*)*o* and *vos*+(*l*)*o* yields the forms *no-lo* and *vo-lo*, but the combination of *lhes*+(*l*)*o* yields the form *lhos*; further, other formal idiosyncrasies involving clusters of clitics cannot be accounted for by pure phonological rules: for example, in a sequence of two adjacent vowels, the first vowel is always absent (*me+o* > *mo*; *te+o* > *to*; *lhe+o* > *lho*);
- (ii) clitic clusters behave syntactically as a single constituent: for instance, the relative order of clitics is identical in proclitic and in enclitic position (e.g. *dá-no-lo* / (*não*) *no-lo* *dás* ‘you (don’t) give it to us’);
- (iii) the ordering of pronominal clitics is fixed; that is, clitics cannot be independently manipulated by movement operations;
- (iv) the relative ordering of pronominal clitics seems to be arbitrary, and it does not appear to depend on syntactic factors (e.g. the dative form precedes the accusative form, contrary to the order of full syntactic phrases);
- (v) there are co-occurrence restrictions among clitics: for example, the impersonal pronoun *se* does not co-occur with accusative forms, but it may appear with dative forms (e.g. Cunha and Cintra 1984: 309) –

see (232a); and the reflexive pronouns do not co-occur with other clitic forms, while other combinations, such as dative and accusative, are possible (see 232b).

- (232) a. **O acordado, deu-se-o ao João.*
 the promised, give-PAST3Sg-IMPS-ACC3Sg to John
Ao João, deu-se-lhe o acordado.
 to John, give-PAST3Sg-IMPS-ACC3Sg the agreed
 ‘As for John, it was given to him what was agreed.’
- b. **As mãos, lavo-mas.*
 the hands, wash-PRES1Sg-REF1Sg-ACC3PlFEM
As maçãs... deu-mas.
 the apples, give-PAST3SgDAT1Sg-ACC3PlFEM
 ‘As for the apples... he gave them to me.’

Clitic clusters in EP thus appear to be inserted as a whole at the moment of phonological instantiation.²⁰⁶

To conclude this section, and according to the proposal made here, we believe to have shown that the phonological idiosyncrasy involving the sequence verb plus clitic is far from being a conclusive argument for the lexical combination of verbs and pronominal clitics: on the one hand, phonological idiosyncrasy is not specific to combinations of words obtained in the lexicon; on the other hand, in a framework such as Hayes’ (1990), it is possible to treat this kind of phenomena as lexical without having to assume that the combination of verbs and clitics is obtained in the lexical component.

4.3. Mesocclisis

Mesocclisis is the third argument put forward in favor of the inflectional status of pronominal clitics, since pronominal clitics appear to precede inflection in this construction (cf. Zwicky 1987a; Halpern 1995). However, a number of facts indicate that the EP construction of future and conditional that yields mesocclisis should not be analyzed as formed in the lexicon, and therefore, mesocclisis is not a valid argument for the lexical combination of verbs and clitics. We will present these facts in the following paragraphs.

It is well-known that future and conditional forms have developed in EP, like in other Romance languages, from an analytic construction found

in Vulgar Latin involving an infinitive verb followed by the present/imperfect form of *habere*.

In syntactic structures that trigger proclisis, instead of enclisis, and in structures without pronominal clitics, it is clear from a phonological point of view that the ancient EP auxiliary *haver* has been reanalyzed as part of inflection: the verb stem and affixes are treated by phonological rules as forming a single phonological word, since there is a single primary stress, and vowel reduction applies to the vowel that would bear stress if reanalysis did not occur (see 233a). In contexts of enclisis, by contrast, the clitic pronoun appears to occur *inside* the verb form, and the whole structure consists of two phonological words: there are two primary stresses, and consequently vowel reduction does not apply to the two stressed vowels (see 233b).

- (233) a. *d[ɐ]r[ɨ]amos* b. *d[ã]r-te-[ɨ]amos* give-COND1Pl(-2SgDAT)
 perceb[ɨ]r[ã]s *perceb[é]r-me-[ã]s* understand-FUT2Sg(-1SgDAT)

As we have seen in chapter 2, in several studies on EP mesocclisis is conceived as involving a verb form inflected for the future or conditional that is split up by the insertion of a pronominal clitic. Specifically, the verb inflected for the future/conditional is (implicitly) seen to form a lexical unit that exists prior to the insertion of pronominal clitics. With differences of implementation, this view is shared, for example, by Mateus (1983), d'Andrade (1988), and Van der Leeuw (1997). In addition, it is usually assumed that the lexical formation of future/conditional has a complex internal structure, and is obtained from the infinitive form of the verb (e.g. Mateus and d'Andrade 2000: 6.2). However, several facts show that clitics are not inserted *inside* the lexically inflected verb, as we will see below.

There are three very frequent verbs in EP where the independent infinitive form does not coincide with the first portion of the future/conditional form (historically related with an irregular infinitive form). This is illustrated in (234).

(234) INFINITIVE	FUTURE	CONDITIONAL	
fazer	<i>fará</i> (* <i>fazerá</i>)	<i>faria</i> (* <i>fazeria</i>)	'to do (3Sg)'
trazer	<i>trará</i> (* <i>trazerá</i>)	<i>traria</i> (* <i>trazeria</i>)	'to bring (3Sg)'
dizer	<i>dirá</i> (* <i>dizerá</i>)	<i>diria</i> (* <i>dizeria</i>)	'to tell (3Sg)'

With mesoclis, both forms may be found (although the one with the regular infinitive is not standard), as in (235a). More interestingly, in the case of verbs morphologically derived from these verbs, there is a clear tendency for speakers to select the *regular* infinitive form for mesoclis, as shown in (235b).

- (235) a. *far-lhe-ia* (?)*fazer-lhe-ia* 'he would do him'
 trar-lhe-ia (?)*trazer-lhe-ia* 'he would bring him'
 dir-lhe-ia (?)*dizer-lhe-ia* 'he would tell him'

b. Without mesoclis

refaria **refazeria*

satisfaria **satisfazeria*

desdiria **desdizeria*

With mesoclis

*?*refar-lhe-ia* *refazer-lhe-ia*

'he would do (him) again'

*?*satisfar-se-ia* *satisfazer-se-ia*

'he would satisfy (himself)'

*?*desdir-te-ia* *desdizer-te-ia*

'he would contradict (you)'

These facts strongly suggest that the clitic is not added to the lexically inflected verb, but rather to the verbal host in the infinitive form.

The material that follows the infinitive verb and the clitic constitutes another problem for the lexical analysis of this construction. According to Zwicky (1987a), in addition to the clitic pronouns, the elements that participate in this construction are the infinitive form plus PN suffixes. The problem with this approach is that the forms that follow the clitic do not match entirely the forms of PN suffixes found in the rest of the verbal paradigm, as shown in (236).²⁰⁷

- (236) a. PN suffixes b. Forms that follow the clitic in mesoclis

		Future	Conditional
1Sg	o/ Ø	–ei	–ia
2Sg	s	–as	–ias
3Sg	Ø	–a	–ia
1Pl	mos	–emos	–iamos
3Pl	m (/ [+nas] /)	–ão (</a[+nas] /)	–iam (</ia[+nas] /)

(based on Mateus et al. 1990)

Both in the future and in the conditional, we can identify the same PN markers as in the rest of the verbal paradigm, listed in (236a), plus some

additional segmental material that, in the case of the future tense, cannot be fully identified with some other inflectional morpheme of EP. This means that what we have after the sequence infinitive verb+clitic is not simply a (PN) inflectional suffix.

In fact, the status of the vowel *a/e* that surfaces in the future tense is not clear under an analysis that treats it as part of a lexical inflectional affix. An alternative analysis of this vowel is to consider it the portion that remains of the stem of the old auxiliary *haver* (from now on, we will refer to this element as *-haver*).²⁰⁸ Under this view, mesoclisism would be formed with the infinitive verb plus a reduced form of the old auxiliary *haver* inflected for the present, in the future tense, and for the imperfect, in the conditional tense. Note that the absence of the stem-vowel in the forms of the conditional, which are marked with the imperfect suffix *-ia*, can be seen to follow from the application of a general rule of the verbal paradigm that is responsible for the deletion of the first of two adjacent vowels in inflectional environments (see chapter 3, section 4 on theme vowel deletion).

There is yet another piece of evidence that can be seen to indicate that the sequence clitic plus *-haver* cannot be treated as a complex of inflectional morphemes. Assuming that the rule mentioned in the preceding paragraph applies to the first of two adjacent vowels in an inflectional environment, this process shows that the sequence clitic plus *-haver* cannot be considered an inflectional complex, nor can it be considered an “inflected pronominal clitic” (Klavans 1985: 116): as illustrated in (237a), the first of two adjacent vowels is deleted in inflectional environment, whereas it cannot be deleted when a clitic+*-haver* sequence is involved (see 237b).²⁰⁹

- (237) a. *falo* (<*fala*+o)
 speak-PRES1Sg
 comia (<*come*+*ia*)
 eat-IMP3Sg
 coma (<*come*+a)
 eat-SUBJ3Sg
- b. *falar-t[j]-emos* (**falar-temos*) ‘(we) will speak to you’
 falar-lh[j]-ei (**falar-lhei*) ‘(I) would speak to him’
 falar-lh[j]-á (**falar-lhá*) ‘(he) will speak to him’

The stress pattern observed in mesoclitic structures is another puzzling fact for an analysis of mesoclisism as a lexical operation consisting in the addition of inflectional suffixes. As shown in chapter 3, section 2, in EP

stress location in the verbal system varies according to tense paradigms: in the present tenses, stress falls on the last vowel of either the root or the stem; and in the past tenses, it falls on the last vowel of the stem. In the future/conditional forms without clitics, the addition of inflection attracts word stress, and there is no secondary stress on the verb root or stem. In all cases, regardless of the exact location of word stress, there is only one stress per word.

As already noted in section 3.2, pronominal clitics do not affect word stress position, although the presence of clitics may yield a sequence of four stressless vowels after the word stress, in a violation of the three syllable stress window. This fact was regarded to show that clitics are not combined with the verb at the moment word stress is assigned in EP. Now, what we find in mesoclisism is that the addition of what is argued to be inflectional suffixation does not cause the primary stress to shift to the right (see 238a), nor is it simply ignored by word stress rules (see 238b). Instead, the element that follows the clitic seems to carry its own primary stress as well (see 238c).

- (238) a. *fal[ɐ]r-te-[é]mos
 b. *fal[á]r-te-[i]mos
 c. fal[á]r-te-[é]mos
 ‘(we) will speak to you’

In other words, the element that follows the clitic behaves phonologically like a word, not as an inflectional morpheme.

The ordering of pronominal clitics with respect to inflectional affixes is also problematic. Based on a survey of a wide number of languages, Bybee (1985) concludes that inflection morphemes are usually ordered following the hierarchy *aspect-tense-mood-person*. If it is assumed that, besides the person/number suffix, a tense morpheme also follows the clitic in mesoclitic constructions (see, for example, Van der Leeuw 1997, and the review in chapter 2, section 4.2), then pronominal clitics should not be able to appear in such a position.

Finally, there is still another fact that seems problematic for an analysis that takes mesoclisism to be a logical result of the inflectional status of pronominal clitics, that can therefore interact with other inflectional affixes (Halpern 1995). This view also presupposes that mesoclisism is a lexical operation of inflection. However, mesoclisism only occurs in the syntactic contexts of enclisis, not of proclisis. Since pre- and postverbal positioning of

clitic pronouns is conditioned by syntagmatic information, as we have seen in section 3.1, mesocclisis, like pronominal cliticization, should not be analyzed as obtained in the lexicon.

We can thus conclude that mesocclisis is not formed in the lexicon, but is rather the result of a syntactic operation.

We will adopt here the hypothesis formulated in Duarte and Matos (2000: 134) in the following terms: “suppose that, in the grammar of standard EP, two forms for the future and conditional co-exist: the ‘new’ synthetic form used in proclisis and enclisis [...], which is inserted fully inflected, and a survival of the analytic form found in Old Romance, where the ancient auxiliary is interpreted as a “lexicalized” T-affix, generated under the T head...”.

In the next paragraphs we will briefly explore this hypothesis.

There are at least two basic syntactic distinctions between the “lexicalized T-affix” and a regular auxiliary (as *ter* ‘to have’): its position relative to the main verb, obligatorily adjacent to the infinitive verb+clitic, and to its right; and its position with respect to pronominal clitics, to their right (see 239a vs. 239b, where the forms of *-haver* and *ter* appear in boldface, and the main verb is underlined).

- (239) a. *O João falar-lhe-á sempre nisso.*
 O João **á-lhe sempre falar nisso.*
 ‘John will always tell him about it.’
 b. *O João **tinha-lhe** já falado nisso.*
 ‘John had already told him about it.’

This behavior and all the bulk of phenomena presented in the preceding paragraphs may be accounted for with the simple assumption that *-haver* is a syntactic, but not a phonological clitic: it has the syntactic distribution of a clitic, but it is not a prosodic clitic, since it is marked with word stress.

The existence of syntactic clitics that are not phonological clitics is not specific to EP, as similar cases can be found in other languages: e.g. disyllabic clitics in Biko (cf. Zwicky 1977), certain particles in Tagalog (cf. Anderson 1992), and the Italian pronoun *loro* (cf. Nespor 1999a).

We therefore propose that *-haver* is treated in the lexical component as a (prosodic) word, and in the syntactic component as an element that is dependent on a verb form, in a way similar to the verbal dependency of pronominal clitics. In other words, we propose that mesocclisis in EP should be analyzed as involving a sequence of syntactic clitics (one of the hypotheses excluded in Zwicky 1987a: 144).²¹⁰

Under this approach, the three constituents that participate in mesoclitic constructions are independently treated in the lexical component and are combined only when syntactic rules have operated. This accounts straightforwardly for the phonological facts associated with mesoclis: the infinitive verb and *-haver* are each prosodized in the lexicon as a prosodic word, and are subject to stress assignment rules, and to vowel reduction; as for the clitic pronouns, they are phonologically deficient and therefore are prosodized in the lexicon as syllables, and undergo vowel reduction.²¹¹

Notice that the analytic future/conditional, unlike the formation of other compound tenses, obligatory involves the incorporation of the main verb with the T-affix, due to the syntactic dependency of the latter element. This operation is thus very similar to the one that derives simple tenses, which also involves verb raising to T. The similarity between the two constructions can be viewed to be at the origin of the morphological reanalysis of the ancient auxiliary as a regular inflectional morpheme in the synthetic forms.²¹²

Let us now consider the details of the syntactic formation of the future/conditional tenses. Our understanding of the formation of the future/conditional is that it is always obtained through the combination of features associated with the infinitive and with the present/imperfect tenses, as in Vulgar Latin. Thus, the syntactic representation is identical both when we find the synthetic and the analytic forms. The choice between one or the other, in turn, is made at the moment of phonological instantiation, in the following way.

The main verb, carrying the infinitive features, raises to the T-head, which is specified with the present/imperfect features. When no clitic pronouns are present, the syntactic representation specified for the adjacent clusters of features that characterize the infinitive and the present/imperfect is formally interpreted by the fully inflected verb. However, when an enclitic pronoun is present, the pronoun attaches to the verb marked for infinitive, in the course of verb raising to T. It is precisely the presence of the clitic to the right of the verb marked for infinitive that blocks the selection of the lexically inflected verb: the clitic intervenes between the infinitive verb and the features in T, and therefore only the analytic form can be selected.

Regardless of the details of implementation, which we will not develop further here, this proposal solves a number of problems that other analyses face:

- (i) it explains the phonological behavior of the elements that participate in the synthetic and in the analytic constructions: the synthetic verb form is obtained in the lexical component through inflectional suffixation, and therefore, the inflected verb is treated as a single prosodic word; in the analytic construction, each element is autonomous in the lexical component, and is prosodized independently, the combination of these elements occurring only postlexically, through an operation of cliticization;
- (ii) it explains the positioning of clitic pronouns in the analytic construction: in contexts of enclisis the pronoun attaches to the right of the infinitive verb, according to the requirements of its lexical entry; the *mesoclitic* effect, in turn, results from the syntactic dependency of *-haver*, which induces the movement, to its left, of the main verb that carries the clitic pronoun;
- (iii) it explains why there are analytic effects only in the contexts of enclisis: it is the presence of the pronoun to the right of the infinitive verb that blocks the insertion of the fully inflected verb, thus revealing superficially the underlying analytic (i.e. syntactic) construction;
- (iv) it is compatible with the existence of contrasts such as those presented in (235): since the analytic construction is obtained via the cliticization of personal pronouns and of *-haver* to a verb inflected for infinitive, while the synthetic form is now obtained through a process of inflectional affixation to a verbal stem, mismatches between the two forms may arise.

To close this topic, we should observe that it is well known that for some (possibly, most) EP speakers there is a different position for pronominal clitics in future/conditional tenses in contexts of enclisis, which is after the whole verb form (e.g. *falarías-lhe* '(you) would speak to him/her'). We interpret these facts to mean that for these speakers the formation of the future/conditional is identical to the formation of simple tenses. In other words, for these speakers the analytic construction is lost. Here, the synthetic fully inflected verb form, which originally developed from a true case of reanalysis of *-haver* as inflection, is always selected. Indeed, this verb form displays the effects we would expect from such a reanalysis: (i) the whole inflected verb behaves phonologically as any other lexically inflected verb, namely, it has only one primary stress; and (ii) most interestingly, the clitic pronoun, being enclitic to a verb form, attaches to the whole inflected verb, thus occurring after it and not inside of it.

It seems to us that this (apparent) shift in the position of pronominal clitics, while simply a logical step under our proposal, is not predicted by an analysis that explains mesocclisis as following from the inflectional status both of clitic pronouns and of *-haver*. This can, therefore, be considered one final argument against the latter position.

To sum up, there is clear evidence that mesocclisis cannot be treated as regular (lexical) inflection, and thus it does not constitute an argument for the lexical attachment of pronominal clitics in EP.

5. Conclusion

To conclude, the facts that have been put forward as evidence in favor of the inflectional status of EP pronominal clitics appear to be either motivated by an erroneous analysis or driven by disputable theoretical assumptions.

We have seen that host selectivity of Romance pronominal clitics may follow from a parameter setting of clitic placement (e.g., pronominal clitics are located with reference to the head of a constituent – the sentence), and thus their similarity with lexical affixes may just be apparent.

We have also proposed to analyze the phonological idiosyncrasy involving pronominal clitics within the framework of precompilation theory, which allows for lexical rules to apply to combinations of items that are not obtained in the lexicon. In this way, we have eliminated phonological idiosyncrasy from the set of compelling arguments in favor of the lexical attachment of pronominal clitics.

Finally, we have shown that mesocclisis cannot be analyzed as a case of inflection operating in the lexicon, and therefore it is not a valid argument for the proposal that pronominal clitics are lexically combined with their host.

Under our account of the facts, EP presents no sound evidence for the inflectional status of pronominal clitics. Since, by contrast, there is clear evidence that the combination of verbs and pronominal clitics cannot be achieved in the lexicon, we conclude that pronominal cliticization should be considered a postlexical operation in this language.

¹⁸⁰ As pointed out in section 2.3.2, Crysmann (1999, 2000a, 2000b) argues that bound pronouns in EP occupy a transitional place between morphology and syntax.

¹⁸¹ The investigation in this chapter was partially presented in Vigário (1998c, 1999a, 1999c, 2000a).

¹⁸² One additional property given in Zwicky and Pullum (1983) is not based on actual empirical observations, but rather follows from theoretical assumptions that we do not share: for Zwicky and Pullum, it is expected that syntactic rules may affect affixed words, but not clitic groups. We believe, in fact, that one of the properties that characterize a syntactic clitic is precisely its special syntactic distribution: on the one hand, they are manipulated by syntax in a fashion similar to independent words, and, on the other hand, they behave like bound elements that are syntactically dependent on another word. Thus, if a syntactic operation affects their host (like verb raising to Comp), it is expected that pronominal clitics are also raised with the verb, since they may cliticize to the verb in a lower position of the syntactic structure.

¹⁸³ In this chapter we use the following abbreviations. NP: nominal phrase; V: verb; N: noun; A: adjective; ADV: adverb; ART: definite article; PRES: indicative present; PAST: indicative past perfect; IMP: indicative past imperfect; SUBJ: subjunctive present; SUBJ IMP: subjunctive imperfect; FUT: future; COND: conditional; INF: infinitive; 1/2/3Sg: 1st/2nd/3rd/person singular; 1/2/3Pl: 1st/2nd/3rd person plural; ACC: accusative pronominal clitic; DAT: dative pronominal clitic; IMPS: impersonal pronominal clitic; REF: reflexive pronominal clitic; MASC: masculine; FEM: feminine; PL: plural; PN: person/number suffix; TMA: tense/mood/aspect suffix; TV: theme vowel. In the pronominal system, gender is indicated only in the feminine forms.

¹⁸⁴ For example, in Cameira (1994) the different accusative clitic forms are obtained from a single underlying form *-lu/*. Consonant loss, in turn, is seen to follow from assimilation and deletion rules. In addition, it is suggested that these rules apply within the clitic group. In Barbosa (1996) this type of data is also assumed to be explained by the application of phonological rules, and is considered to constitute phonological evidence for the different directions of phonological cliticization of proclitics and enclitics (see section 2.3.2 for a review).

¹⁸⁵ Although lexical affixes typically have fixed positions, i.e. they either attach to the left or to the right of their bases, there are at least two languages that have been claimed to have mobile affixes: Afar, a East Cushitic language, and Huave, a language of Oaxaca-Mexico (cf. Fulmer 1991 and Noyer 1994, respectively). Nevertheless, in none of these cases is affix location argued to be dependent on phrasal information, unlike in the case of EP pronominal clitics, as we will see in the next paragraph.

¹⁸⁶ The relevant domain is tentatively defined in Frota and Vigário (1996) as the syntactic domain CP and the intonational phrase (I) prosodic domain. That is, the

items that cause proclisis must be both within the same CP and within the same I as the clitic.

¹⁸⁷ As a matter of fact, enclisis in contexts of proclisis is very frequent in younger generations, i.e. there is a recent regression of proclisis, as reported in Frota (1994) and Duarte, Matos, and Faria (1995). This shift in clitic placement follows naturally under the view presented therein that in this new grammar pronominal clitics are a step further towards the affixal status. Note that all the facts we are describing concern exclusively the (preceding) stage where proclisis and enclisis co-exist, and no claim is made here as to the reanalysis of pronominal clitics as inflectional affixes in future stages of EP.

¹⁸⁸ We should add that, although interpolation is distributionally very limited (nowadays it is only found with negation, and possibly with a few other simplex adverbs), and it only occurs in certain styles and/or dialects, cases of interpolation can still arise in *lapsus linguae*, as documented in Frota (1994). This is suggestive of the psychological reality of a certain *mobility* of clitic pronouns in EP.

¹⁸⁹ We thank Danièle Godard for drawing our attention to this point.

¹⁹⁰ For the sake of comparison, all the examples show the same underlying vowel followed by the nasal segment (/e[+nas]/).

¹⁹¹ It should be noticed that there are a few cases involving compound structures which seem to have been reanalyzed as simple words so that final diphthongization resulted in simple nasalization of the preceding nucleus and deletion of the nasal segment, as in word internal position:

- (i) *sensaboria* [ẽ]/??[ẽ] (cf. *sem* sabor [ẽ] ‘without flavor’)
 ‘tediousness’
 sempre-em-pé [ẽ]/??[ẽ] (cf. *em* pé [ẽ] ‘on foot’)
 lit. always on foot

This is, in fact, what we would expect in verb final position if pronominal clitics had been reinterpreted as internal to the verb form.

¹⁹² Note that mesoclisism is not a counter-argument to this generalization, under the view presented in section 4.3.3.

¹⁹³ This argument builds on a similar observation on the distinction between inflection and derivation, made in Anderson (1992: 79). On the notion of *portmanteau* see also, for example, Bauer (1987: 18).

¹⁹⁴ As pointed out in Pesetsky (1985), Sproat (1988), and Booij and Lieber (1993), among many others, the existence of forms like *unhappier* give rise to a paradox: while from a phonological point of view, the word must be derived from *happier* (since *unhappy* would constitute a trisyllabic base), from a morphological point of view it must be derived from *unhappy*, given its semantic interpretation. See section 1.2.2 for a discussion on the resolution of such paradoxes.

¹⁹⁵ It should be noticed, nevertheless, that there are many disyllabic bases to which *-ez* is added. Evaluative suffixes seem to exhibit a similar behavior, according to

Skorge (1956), cited in Villalva (1994: 5.3.3): an evaluative form like *-inho* or *-ito* tends to select disyllabic bases, while z-evaluative suffixes like *-zinho* or *-zito* tend to select polysyllabic bases. As Villalva observes, there is, nevertheless, great variation in this case, as other factors also seem to play a role in the presence of one type of suffix or the other.

¹⁹⁶ Crysmann (2000b) suggests that the fact that bound pronominals do not attach to participles is an instance of an “arbitrary gap”. We believe, on the contrary, that this *gap* may result from the properties of the participle itself instead of the clitic’s, as it is well-known that participles do not behave like full verbs in many respects.

¹⁹⁷ In other syntactic accounts of cliticization, such as Sportiche (1992), pronominal cliticization is treated as inflection (see, for example, Manzini 1998 for a review of syntactic approaches to cliticization). Spencer (1991: 391–392) observes that syntactic treatments where *clitics* and inflection are conceived as abstract features associated to functional categories imply that it is necessary to find other, non-syntactic, criteria for distinguishing “full words from clitics from affixes from other morphological phenomena”.

¹⁹⁸ Besides the syntactic information that plays a role in pronominal clitics distribution, EP clitic positioning is shown in Frota and Vigário (1996) to be sensitive to prosodic information. They suggest that proclisis is triggered by heavy function words that have to precede the clitic pronoun within the same CP and the same intonational phrase. Following this hypothesis, Duarte and Matos (2000) conclude that proclisis in EP is in general an instance of “*Move* occurring between Spell-Out and the P-A Interface”. Note that, independent of the exact point in the grammar where clitic position is obtained, it still has to precede lexical insertion, as the form of clitics varies according to their position with respect to the verb (e.g. *come-lo* ‘(you) eat it’ vs. *não o comes* ‘(you) don’t eat it’).

¹⁹⁹ Each of these forms may also appear in the feminine (*a/la/na*) and/or in the plural (*os, as/los, las/nos, nas*). For the sake of simplicity, in general we omit them, since they behave (*mutatis mutandis*) like the masculine singular form.

²⁰⁰ In normative grammars no distinction is made between nasal diphthongs found in the 3rd person plural and nasal diphthongs found in the 3rd person singular: in both cases the form of the clitic is said to be *no* (e.g. Cunha and Cintra 1984: 280).

²⁰¹ There is another instance of consonant deletion, affecting words formed with the evaluative suffix *-zinho*: when the base ends in fricative, this consonant is obligatorily deleted. Thus, the diminutive form of *azuis* ‘blue-Pl’ is *azuizinhos* and not *azuiszinhos*. The first of two fricatives also tends to be deleted within the word (e.g. *transgênico* ‘transgenic’). Across words, on the contrary, the two fricatives may occur, as in *as zebras* ‘the zebras’, where [ʒz] or [ʒ] are possible realizations, but not [z]. We can add that the latter possibility, but not the penultimate one, may be found in more conservative dialects.

²⁰² Notice that the lexical processes involving clitics and their hosts in Polish, referred above in section 4.2.2, can also be an instance of precompiled phrasal phonology (cf. Halpern 1991).

²⁰³ Cases of allomorphic alternations that are accounted for in a similar fashion are the alternation in the form of the English indefinite article *a/an* and in the form of Spanish feminine article *la/el* (cf. Hayes 1990).

²⁰⁴ For those speakers that do not differentiate between verb forms ending in nasal diphthongs, the relevant context for the insertion of the allomorph *no* should be [... [+nas]/_{vb} _], instead. Actually, there are other instances of variation related to the selection of pronominal clitic allomorphs in EP. To mention just two examples, in some dialects, the allomorph *lo* is also selected in nasal contexts; *o*, instead of *lo*, may also be selected when the verb ends with a fricative, and in some cases, a linking [j] may appear between the clitic and a preceding fricative (cf. Mota 2000).

²⁰⁵ Notice that the relevant information assessed by these rules is exclusively morphosyntactic and segmental, not prosodic. Therefore, the application of these processes cannot support any claim related to the direction of phonological cliticization of pronominal clitics, as suggested in Barbosa (1996) (see section 2.3.2). In fact, nothing prevents the occurrence of precompiled phrasal rules between two elements that are not under the same prosodic node.

²⁰⁶ Certain combinations of prepositions and articles in EP, like in Italian (cf. Napoli and Nevis 1987) or French (cf. Zwicky 1987b), seem to be analyzable in a parallel way.

²⁰⁷ In the Perfect tense, TMA and PN are portmanteau forms (e.g. Villalva 1994) and therefore these are not considered in (37a).

²⁰⁸ The existence of two forms that alternate (a/e) fits well with the hypothesis that this element has a stem status. Indeed, while there are no cases of formal alternations involving inflectional affixes, there are several instances of stem alternations in EP within a single conjugation (e.g. *valho/vales* '(I/you) worth'; *trago/trazes* '(I/you) bring'; *faço/fazes* '(I/you) do'—e.g. Mateus 1975: 2.5).

²⁰⁹ Notice that the same rule could explain some specific phonological forms associated with dative and accusative clitic combinations: *mo(s)/ma(s)* (<me+o(s); me+a(s)); *to(s)/ta(s)* (<te+o(s); te+a(s)); *lho(s)/lha(s)* (<lhe+o(s); lhe+a(s)). Thus, clitic clusters, but not the sequence clitic plus *-haver*, show the behavior expected from inflected pronouns. Notice further that, according to some analysis, vowel deletion only applies in the present tenses, since in the past tenses, the absence of vowel is seen to follow from a rule of fusion (cf. section 3.3). In any event, we can establish the generalization that whenever two adjacent vowels appear in verbal inflectional environments, some process applies that prevents the adjacent vowels to surface. Therefore, also under this formulation of the facts, clitics do not behave like inflectional affixes.

²¹⁰ We use the term *syntactic clitic* to express the fact that *–haver* has a syntactic distribution different from other *free* auxiliaries, as it is *dependent* on the infinitive verb+clitic sequence.

²¹¹ We follow the view presented in section 1.2.2, and argued in Booij (1988), Booij and Lieber (1993), and Nespor (1990), that prosodic structure up to the prosodic word level is built in the lexicon. See, in addition, section 5.3.2 for the arguments in favor of the lexical prosodization of clitics, and sections 3.1 and 3.2 respectively for the lexical status of stress assignment and vowel reduction in EP.

²¹² Reanalysis has presumably given raise to a portmanteau suffix. That is, the adjacent morphemes that realize the infinitive and the present/imperfect features were probably reanalyzed as a single unit. Thus, the two morphosyntactic sets of features are realized in a single suffix. In fact, in EP there are no instances of two co-existing TMA morphemes, thus suggesting that in this language there is only one slot for TMA markers in the inflectional template. Notice that, as we will see below, reanalysis is only possible in the synthetic form (the one without an intervening clitic), since in the analytic form, the two sets of morphosyntactic features are realized by two distinct formal units. See Villalva (1994: 4.3.3.2) for a proposal on the morphosyntactic features associated with tense, mood, and aspect in EP, and for a classification of portmanteau suffixes in the verbal system. Notice that the formation of future and conditional is excluded from Villalva's observations: because of the possible occurrence of internal clitics, these are viewed as compound forms (cf. Villalva 1994: chap.4, note 57).

Chapter 5

The prosodic word

1. Introduction

As we have seen in chapter 1, there are two major issues concerning the definition of prosodic domains crosslinguistically: on the one hand, languages differ in the type of phenomena that either refer to or characterize a given prosodic domain; on the other hand, languages vary in the way certain morphosyntactic units are integrated into prosodic structure. The present chapter focuses on these two topics, as far as the prosodic word domain is concerned. We begin with a review of the phonological phenomena that may identify the prosodic word in EP (section 2). We proceed in the ensuing sections with the observation of the behavior of a number of constructions with respect to the phenomena shown to characterize the prosodic word domain. This investigation will enable us to establish the prosodic structure associated with each construction (sections 3 and 4).

The presence of a single word stress constitutes the common feature of the constructions analyzed, which comprise derived words, inflected words with affixes that do not form stress domains independent of their base, as well as word plus clitic combinations. The similarity between affixes and clitics in terms of their prosodization emerges from the presentation in sections 3 and 4, whereas the phonological differences between these units are inspected in section 5. Finally, we conclude this chapter with a discussion on the definition of the prosodic word domain in European Portuguese.

2. Phonological properties of prosodic words

As we saw in chapter 1, section 6.1, the phenomena summarized in (240) are amongst the most common diagnostics for the prosodic word in different languages.

(240) Common diagnostics for the prosodic word

- word stress assignment and/or location
- (other) phonological phenomena that refer to the prosodic word domain
- phonotactic generalizations
- deletion under identity
- clipping
- minimal word requirements

The discussion below will show that not all these phenomena yield reliable diagnostics for the prosodic word domain in EP.²¹³

First of all, some of these phenomena only play a role in the lexical phonology of the language. This is the case of word stress location, since word stress is operative only at the lexical level (see chapter 3, section 2)²¹⁴, of a number of phonological processes described in chapter 3 that refer to the prosodic word, and that are active only in the lexical component, as well as of phonotactic generalizations, which also seem to hold only at the lexical level, as we will see below.

Other phenomena, like deletion under identity and clipping (see chapter 1, section 6.1), necessarily involve constructions with more than one word stress in EP, and thus with more than one prosodic word. For this reason, we will not consider them here. Instead, the relevance of these processes as diagnostics for prosodic wordhood in this language will be discussed in chapter 6, section 4.

Finally, like Brazilian Portuguese and Irish and unlike many languages cited in chapter 1, section 6.1, EP does not seem to be sensitive to minimal word requirements. Indeed, as previously noticed in Bisol (2000), in Portuguese there are monosyllabic stressed words that are formed by light syllables, like those presented in (241). This means that neither the size of words nor their internal syllable make-up may be used to determine whether or not a given string may form a prosodic word.

(241)	<i>ri</i>	[rí]	‘laugh’		<i>nu</i>	[nú]	‘naked’		
	<i>dê</i>	[dê]	‘give’		<i>dou</i>	[dó]	‘(I) give’		
	<i>pé</i>	[pɛ́]	‘foot’	<i>pá</i>	[pa]	‘spade’	<i>nó</i>	[nɔ̃]	‘knot’

Besides this, it appears that in this language (at least in the variety studied here) there are no phonological processes that result in the avoidance of monosyllabic words ending with a vowel.²¹⁵ By contrast, phono-

logical processes occur that may actually cause either monosyllabic words ending with a glide or consonant to surface only with a vowel, or disyllabic words to become monosyllabic: for example, /ow/ monophthongization (cf. Mateus 1975: 42), syllable degemination (see chapter 3), and resyllabification (see chapters 3 and 8, and the examples in 4 further below).²¹⁶

We conclude, therefore, that the prosodic word in EP is not crucially affected by minimal word requirements. It is, nevertheless, important to observe that monosyllabic words are relatively rare in the language. Indeed, an inspection of the corpus of the most frequent words in EP organized by Bacelar, Marques, and Segura da Cruz. (1987) shows that among the ca. 7000 (inflected) forms listed, only 138 are monosyllabic, and of these only 28 end in a vowel.²¹⁷ EP is therefore similar in this respect to Italian, where, according to a dictionary search by Thornton (1996), monosyllabic nouns represent less than 1% of the words of the language.²¹⁸

Let us now turn to the tests that we believe to confidently cue the prosodic word in EP. These include phonotactic generalizations, the presence of word stress and the application of processes that refer to the prosodic word domain. In each case, however, it is necessary to identify the point in the grammar where these phenomena are operative. We will consider each in turn.

The prosodic word is a domain for phonotactic restrictions. In particular, the prosodic word initial position imposes restrictions to the occurrence of certain segmental material, since not all segments allowed to begin a syllable in word internal position may occur word initially. These segments are listed in (242), together with illustrative examples showing the same segments starting a syllable in prosodic word internal position.

(242) * _ω [_	:	[ʎ]	versus	<i>alho</i>	‘garlic’
			[ɲ]		<i>ganho</i>	‘(I) win’
			[r]		<i>caro</i>	‘expensive’
			[i]		<i>pueril</i> ²¹⁹	‘childish’

Notice that among these segments a distinction should, nevertheless, be made, since the impossibility of prosodic word initial [r] and [i], but not of [ʎ] and [ɲ], is the result of the (non-/)application of phonological processes, namely, the application of /r/-strengthening in prosodic word initial position and the absence of (full) vowel reduction in prosodic word initial position (see chapter 3, sections 9 and 10, respectively).

The fact that a clitic word such as the dative pronoun *lhe* starts with [ʎ] constitutes evidence that the notion of *word* relevant to phonotactic generalizations is phonological rather than morphological (see section 4.1 on the phonological characterization of clitics).

It is clear, furthermore, that these generalizations hold only at the lexical level, since prosodic word final vowel deletion (see chapter 3, sections 12–14) and/or resyllabification may result in prosodic words starting with [ʎ], [ɲ] or [r], as shown by the examples in (243), and the postlexical prosodization of the clitic *lhe* in preverbal position yields a prosodic word starting with [ʎ] (see section 4.3.2 on the prosodization of proclitics).

(243) Lexical prosodic words		Postlexical readjusted prosodic words ²²⁰
(malha) _ω (original) _ω	(ma) _ω ([ʎ]original) _ω	‘original mesh’
(tenho) _ω (ainda) _ω	(te) _ω ([ɲ]ainda) _ω	‘(I) still have’
(ver) _ω (amigos) _ω	(ve) _ω ([r]amigos) _ω	‘(to) see friends’

Similarly, according to d’Andrade and Rodrigues (1999), a schwa may optionally occur at the beginning of a word starting with a coronal fricative followed by a consonant (e.g. *escola* ‘school’ [ʃkɔlɐ] or, quite marginally we believe, [iʃkɔlɐ]). It can be assumed that this possibility results from a more general postlexical process of insertion of [i] to fill empty syllabic nuclei created by the syllabification algorithms (see Mateus and d’Andrade 2000: chap.3). This is consistent with d’Andrade and Rodrigues’ (1999) proposal that words with an underlying initial sequence comprising an underspecified fricative followed by a consonant (that the authors represent as /sC/) are syllabified with an initial empty V-slot. The postlexical nature of this process may be inferred, for instance, from its sensitivity to speech rate.²²¹ Thus, unlike in other languages, such as Dutch (cf. Booij 1995, 1999), and German (cf. Wiese 1996; Hall 1999a), schwa appears to be allowed (although marginally) at the beginning of a prosodic word, postlexically.²²² This again suggests that phonotactic generalizations concerning the prosodic word domain operate only at the lexical level in EP.

Pointing in the same direction is the impossibility of [ʎ] to be preceded by any consonant within a word at the lexical level. Postlexically, however, the sequence C-[ʎ] may arise as a consequence of the combination of the enclitic *lhe* with a verb that ends with a consonant (as in *dar-lhe* ‘give-him’ or *dás-lhe* ‘(you) give-him’).²²³ Notice that, since it will be shown in

section 4.3.1 that enclitics are incorporated into the prosodic word that dominates the verb, the prosodic structure associated with the verb-clitic combination may not be invoked to explain the observed lexical/postlexical asymmetry.

Finally, although [R] cannot be preceded by [r] prosodic word internally (see chapter 3, section 9), the sequence [r]-[R] may occur in a string obtained postlexically, after the concatenation of words (cf. *mar revolto* ‘troubled sea’).

The presence of word stress is also a cue for prosodic wordhood, since it can be assumed that there is a one-to-one correspondence between a (word) stressed unit and a prosodic word, as already stated in (28) in chapter 1, repeated below under (244).

- (244) A prosodic word must be assigned one and only one (word) primary stress

Since word stress has the properties of a lexical phenomenon and thus is not operative postlexically (see chapter 3, section 2), we should point out, however, that stress location may not be used to assess the *limits* of the prosodic word domain (see also section 4.3.1).

The presence of word stress is correlated with the occurrence of other phonological phenomena that characterize stressed positions, such as pitch accent assignment and focal stress (see chapter 3, sections 18 and 19, respectively). These phonological events thus provide additional support for the prosodic word status of a given constituent.

Conversely, the presence of word stress implies the impossibility of occurrence of phonological processes that apply in unstressed environments. This is an important indication of the prosodic word status of a given unit, in particular in the case monosyllabic words. Among the processes that may not affect stressed segments are vowel reduction, as exemplified in (245) (see also chapter 3, section 3), and the semivocalization or deletion of the first of two adjacent vowels, as illustrated in (246) (see also chapter 3, sections 11.1 and 12 to 14, respectively). Thus, if any of these processes apply to a monosyllabic word, this constitutes evidence that the relevant word is not stressed and therefore lacks prosodic word status (see section 4.1, on the characterization of clitics).

We summarize in Table 2 the phonological phenomena that refer to the prosodic word domain in EP, together with the indication of the locus in the grammar where they are operative, according to the proposals in chapter 3.²²⁴

In the next sections we use the phenomena listed above as diagnostics for the prosodic word domain and for the prosodic structure assigned to each construction involving a single primary word stress.

Table 2. Phonological phenomena that refer to the prosodic word domain and the locus in the grammar where they are operative.

Phenomena that refer to the Prosodic Word domain	Locus of operation
Heterosyllabic /e/-centralization	Lexical level
Glide insertion to break a hiatus	
Lowering of stressless vowels in final syllables closed by /r/	
Initial /r/-strengthening	
Initial vowel realization (neutralization)	
Initial vowel realization (feature specification)	Postlexical level
V2 semivocalization	
Final non-back vowel deletion	
Final round vowel deletion	
Final central vowel deletion	
Syllable degemination	
Initial stress	
Emphatic stress	
I-initial H-tone assignment	

3. Derived and inflected words corresponding to a single stress domain

Words with affixes that do not constitute stress domains independent of their bases may be differently prosodized in EP depending on the position of the affix relative to the morphological base. This will be shown in sections 3.1 and 3.2, where the prosodization of suffixes and prefixes, respectively, is studied.

3.1 Prosodization of suffixes

Except for a very limited number of suffixes that select a morphological word as their base, instead of a stem or a theme (cf. Villalva 1994 and section 6 below for the definition of these terms), which will be treated in chapter 6, in the large majority of suffixed words the base and the suffix form together a single prosodic word.²²⁵ We will proceed with the demonstration of the facts that support the previous statement, thus revealing a behavior that will contrast with the one found in prefixed words.

The examples in (247) show that suffixed words constitute a single domain for word stress.

- | | | | | |
|----------|------------------|----------------|------------------|---------------|
| (247) a. | <i>Arma</i> | [árɐ̃] | <i>sOnho</i> | [sóɲu] |
| | ‘arm’ | | ‘dream’ | |
| b. | <i>armAr</i> | [ɐ̃ɐ̃már] | <i>sonharÁs</i> | [sɔɲɐ̃ɐ̃ɐ̃ɐ̃] |
| | ‘to arm’ | | ‘you will dream’ | |
| c. | <i>armamENto</i> | [ɐ̃ɐ̃mɐ̃mɐ̃tɔ] | <i>sonhIto</i> | [sɔɲítu] |
| | ‘armament’ | | ‘little-dream’ | |

Indeed, in the forms (247b–c) the absence of word stress associated to the final vowel of the base suffixes attach to, which correlates with the presence of vowel reduction, shows that this morphological constituent does not form a prosodic word independent of the suffixes.

The impossibility of prosodic word final vowel deletion further shows the lack of prosodic word status of the base (word final vowel deletion phenomena are described in chapter 3, sections 12 to 14). This is illustrated in (248), where the morphological division is given inside parentheses.

- | | | | | |
|-------|---------------------|------------------|------------|--------------|
| (248) | <i>continuidade</i> | (continu -idade) | [u]/[w]/*0 | ‘continuity’ |
| | <i>ingenuidade</i> | (ingenu -idade) | [u]/[w]/*0 | ‘ingenuity’ |
| | <i>crueza</i> | (cru -eza) | [u]/[w]/*0 | ‘crudity’ |

The absence of independent prosodic word status of these affixes, in turn, is supported by the non-application of prosodic word initial /r/-strengthening to suffixes starting with [r] (see chapter 3, section 9), and by the impossibility of emphatic stress at the left edge of the suffix (see chapter 3, section 17). These facts are illustrated in (249).²²⁶

- | | | | | |
|----------|-------------------------|----------|------------------------|------------------------------|
| (249) a. | <i>pediria</i> | [r]/*[R] | (ped i - ria) | ‘(he) would ask’ |
| | <i>saberemos</i> | [r]/*[R] | (sab e - re - mos) | ‘(we) will know’ |
| b. | * <i>marisQUEIrinha</i> | | (marisc - eir - inh a) | ‘little sea food restaurant’ |
| | * <i>amaBilissimo</i> | | (am a - bil - issim o) | ‘very kind’ |

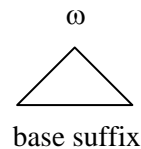
This prosodization is further supported by the realization of the vowels in the suffix initial position. As we saw in chapter 3, section 10, non-central non-high stressless vowels exhibit some variability in height when they occur in word initial position, and they are not subject to (full) vowel

reduction. However, this sort of variability does not occur with vowel initial suffixes, which behave like prosodic word internal vowels for the purposes of vowel reduction, as shown in (250).²²⁷

- (250) *belEza* [é]; *beleZUra* (bel -ez -ur a) [i]/*[e]/*[i]
 ‘beautiful’ ‘very beautiful’
venenOso [ó]; *venenOsÍssimo* (venen -os -íssim o) [u]/*[o]/*[ɔ]
 ‘toxic’ ‘very toxic’

The facts reported in the preceding paragraphs show that there is no prosodic word boundary between the morphological base and the suffix: on the one hand, the segments at the right edge of the morphological base do not behave as prosodic word final; on the other hand, the segments that start suffixes do not behave as prosodic word initial. Consequently, the only prosodic structure that is compatible with the data is the one in (251).

- (251) The prosodization of suffixes in EP



According to the proposed structure, suffixes are incorporated into the prosodic word that dominates their base, and thus suffixed words behave *phonologically* like simple words.

3.2 Prosodization of prefixes

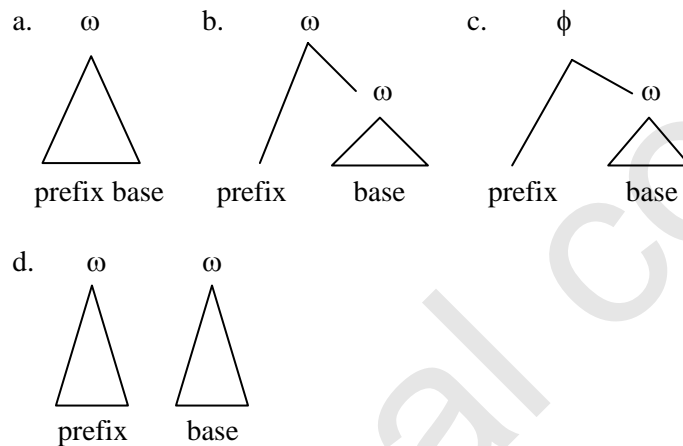
As mentioned in chapter 1, the prosodic status of prefixes varies cross-linguistically. Nevertheless, in all languages unproductive or non-transparent prefixation often yield prosodic structures similar to simple words (e.g. Nespor and Vogel 1986; Peperkamp 1997a), even in languages like English or French, where (monosyllabic) prefixes may form prosodic words independent of their bases (cf. Raffelsiefen 1999a and Hannahs 1995b, respectively). In order to isolate this variable, we will first consider transparent constructions with two unstressed prefixes that we believe are the most productive stressless prefixes of EP: *re-* (‘re-’) and *des-* (‘un-’). Indeed, these prefixes (i) allow for the formation of new words, (ii) they

often enter in constructions with compositional meaning, (iii) they often appear attached to morphological bases that correspond to existing words, and (iv) they present regular phonological behavior.

We close this section with a comparison between the phonological behavior associated with stressless prefixes in transparent constructions and the behavior of prefixes in non-transparent constructions.

According to recent proposals reviewed in chapter 1, section 6, there are at least four theoretical possibilities for the prosodization of prefixes, as depicted in (252) (e.g. Peperkamp 1997a; Raffelsiefen 1999a; Hall 1999a).

(252) Possible prosodizations of prefixes



As for the configuration in (252d), it cannot be associated with this type of prefixation because *re-* and *des-* lack word stress and thus may not form prosodic words independent of their hosts. This assertion is supported by the fact that these prefixes have [i] as their only vowel (see 253a), and by the possibility of semivocalization of the prefix vowel when the base it attaches to starts with a vowel (see 253b).

- (253) a. *refazer* [i] 'to make once again'
 desfazer [i] 'to undo'
 b. *readaptar* [j] 'to readjust'
 realojar [j] 'to accommodate again'

Having ruled out the possibility that these prefixes may form independent prosodic words, three other prosodic configurations are still available: prefixes may be *incorporated* into the prosodic word that includes their morphological base, as in (252a), they may be *adjoined* to it, as in (252b), or they may be directly attached to ϕ , as in (252c). A number of arguments indicate that these prefixes are *adjoined* to the prosodic word that contains their morphological base. Indeed, as we will see in the next paragraphs, not only the segments belonging to the prefix show the phonological behavior of prosodic word initial elements, but also base initial segments pattern like segments at the beginning of the prosodic word.

Let us start by showing that the left-edge of the base corresponds to a prosodic word initial position. The phonological processes that we have identified in chapter 3 to occur in prosodic word initial position also characterize the base left-edge. These include initial /r/-strengthening (see 254a), the absence of full vowel reduction in initial position and the variability in the height features that characterizes initial vowels, but not word internal ones (see 254b versus 254b'), and emphatic stress assignment (see 254c).²²⁸

- (254) a. *rerubricar* [R] 'to sign again'
 reregionalizar [R] 'to regionalize again'
- b. *reorganizAr* [o]/[ɔ]/*[u] b'. (gOlo - [o]) g[u]lInho
 'to reorganize' 'goal' 'little goal'
 desocupAr [o]/[ɔ]/*[u] (desOla - [ɔ]) des[u]lAr
 'to vacate' '(he) distresses' 'to distress'
 reeditAr [e]/[i]/*[i] (diEta - [ɛ]) di[i]tÉtico
 'to publish again' 'diet' 'dietetic'
 deseducAdo [e]/[i]/*[i] (desErto - [ɛ]) des[i]rtAr
 'impolite' 'desert' 'to desert'
- c. *reINSpeccionAr* 'to reexamine'
 desFLOrestaçãO 'deforestation'

That these prefixes are not incorporated into the prosodic word that dominates the base is further supported by the non-application of heterosyllabic /e/-centralization, a process that applies within the prosodic word domain when /e/ is followed by a heterosyllabic palatal segment (see chapter 3, section 6). This is shown by the contrast between prefixed and non-prefixed words given in (255a) and (255b), respectively.²²⁹

- (255) a. *reisolar* *[ɐ] ‘to isolate again’
 reilustrar *[ɐ] ‘to illustrate again’
 b. *veicular* [ɐ] ‘to diffuse’

In fact, if the process applies *within* the prosodic word domain, the relevant segments must be dominated by the prosodic word node. Thus, under the definition of dominance of Chomsky (1986: 7) provided in (256) and given the assumption that the prefix is adjunct to the prosodic word that dominates its morphological base, the configuration for the rule to apply does not obtain in the prefixed words in (255) (see also Booij 1996a and Peperkamp 1997a on the importance of the notion of *dominance* for the application of rules that operate within prosodic domains).

- (256) α is dominated by β if it is dominated by every segment of β .²³⁰

Finally, these prefixes also behave like prosodic word initial: they undergo initial /r/-strengthening, as illustrated in (257a), and they may bear emphatic stress, as shown in (257b). These processes affect prosodic word initial segments regardless of their position within the phonological phrase (see 257a’ and 257b’).

- (257) a. *rerubricar* [R] ‘to sign again’
 reregionalizar [R] ‘to regionalize again’
 a’. (vamos *rerubricar*)_φ [R] ‘let’s sign again’
 (vamos *reregionalizar*)_φ [R] ‘let’s regionalize again’
 b. *REinspeccionar* ‘to reexamine’
 DESflorestação ‘deforestation’
 b’. (vamos *REinspeccionar*)_φ ‘let’s reexamine’
 (imensa *DESflorestação*)_φ ‘huge deforestation’

With respect to the possibilities of assigning emphatic stress in prefixed words, we can add external data in order to support our description. In an informally collected corpus presented in Appendix I, based on the observation of the speech of politicians, TV news reporters, and lecturers, the examples we found where emphatic stress is assigned to the second syllable of the word (from left to right) usually involve prefixed words.²³¹ Besides this possibility, in most of the examples the emphatic stress is assigned to

the first syllable, which corresponds to the prefix, regardless of its position within ϕ . The forms in (258) were extracted from this corpus.

- | | |
|-------------------------------------|-------------------------------|
| (258) Emphatic stress on the prefix | Emphatic stress on the base |
| <i><u>I</u>legalidade</i> | <i>i<u>NA</u>creditável</i> |
| ‘illegality’ | ‘incredible’ |
| <i><u>RE</u>distribuí-las</i> | <i>re<u>CLA</u>ssificação</i> |
| to redistribute-it-Pl-FEM | ‘reclassification’ |

We should add, to conclude, that syllable degemination may also apply when the second syllable involved belongs to a prefix, as shown in (259).

- (259) *pOde des-sintonizAr* [didiʃ]/[diʃ] ‘(he) may untune’
impEde desflorestações [didiʃ]/[diʃ] ‘(he) prevents deforestations’

According to the formulation of the process presented in chapter 3, section 15, this could constitute additional evidence for the prosodization of prefixes as prosodic word initial. Nevertheless, as it will become clear in the following sections and in chapter 6, it is possible to formulate the rule as applying simply to syllables in prosodic word final position, within an intonational phrase, without further specifications as to the prosodic domain the following syllable belongs to (see 260).

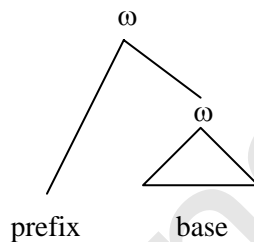
- (260) Syllable degemination (alternative to the formulation in 3.14)

- a. $V_1 \rightarrow \emptyset / [\dots [\dots C_a _]_\omega C_a V_2 \dots]_I^{\max}$
 b. $C_a \rightarrow \emptyset / [\dots [\dots _]_\omega C_a \dots]_I$

(where the syllables headed by V_1 and V_2 conform to the conditions stated in 190 of chapter 3)

This formulation implies that the segments to be deleted occur necessarily in prosodic word final position, but the following syllable may, for example, belong to a clitic that may be attached to a higher-level prosodic domain, in a configuration where it is not prosodic word initial.²³² Thus, syllable degemination does not constitute an unquestionable diagnostic for prosodic word initial position.

(261) The prosodization of prefixes in EP



In prefixed words like those in (262), the vowel that starts the morphological base undergoes full vowel reduction, unlike what we have seen to usually happen with *re-* and *des-*.

- (262) a. *odOr* [o]/[ɔ]/*[u] *inodOr* [u]/*[o]/*[ɔ]
 ‘smell’ ‘odorless’

b. <i>eficAz</i>	[e]/[i]/*[i]	<i>ineficAz</i>	[i]/*[e]/*[i]
‘efficient’		‘inefficient’	
<i>evitÁvel</i>	[e]/[i]/*[i]	<i>inevitÁvel</i>	[i]/*[e]/*[i]
‘avoidable’		‘unavoidable’	
c. (<i>odorizANte</i>)	[o]/([ɔ])/*[u]	<i>desodorizANte</i>	[u]/[o]/([ɔ])
‘odoriferous’		‘deodorant’	

Some of these forms deserve a brief comment, though. As for (262a), we should remark that the form that appears in the dictionary is *inodoro*, while *odoro* does not exist as an independent word. Nevertheless, the form *inodor* is also used. In both cases, the vowel that starts the morphological base of the prefixed word may only be realized as high. As for the example in (262c), we should observe that the form *odorizante* is a *possible* but not an *actual* word in EP, and the word *desodorizante* may have two meanings: one that is compositional (a product that eliminates (bad) smell); and one that is non-compositional, which is in fact nearly the opposite of its compositional meaning (a product that adds (good) smell).

All the forms in (262) show that the incorporation of the prefix has occurred, since when no incorporation occurs the vowels that begin the morphological base *cannot* be fully reduced. As for the cases that allow for alternative realizations (e.g. 262c), it is possible that two analyses of the words are available, one compositional and the other reanalyzed as a simple word. In addition, when the prefix involved is a productive one it is plausible that the complex morphological nature of the relevant words is more accessible than when the prefix is not (so) productive – as in the contrast between *in-* and *des-*.

That these words involve incorporation rather than adjunction of the prefix to the prosodic word that dominates its base is further supported by the impossibility of emphatic stress on the second vowel of the word when the prefix is *in-*, as shown in (263).²³³

- (263) *Inodor* / ^x*iN**Odor* ‘odorless’
Ineficaz / ^x*iNE**ficaz* ‘inefficient’
Inevitável / ^x*iNE**vitável* ‘unavoidable’

Other instances of non-transparent prefixed words that behave phonologically like simple words are those that involve prefixes which are stressed in their productive transparent use, such as *prÉ-* [prɛ] ‘pre-’, *prÓ-* [prɔ] ‘for’, and *pÓs-* [pɔʃ] ‘post-’ (see chapter 6 for the prosodization of

stressed prefixes). The forms in (264) have a non-compositional meaning; the prefix is stressless, as shown by the fact that its vowel obligatorily undergoes vowel reduction; and the morphological category of the base is not necessarily the same as the one selected in productive (stressed) counterparts (as in the case of *pró-* and *pós-*, which attach productively to adjectival but not verbal bases).

- (264) *prever* [pri]/*[prɛ] 'to foresee'
pospor [puʃ]/*[pɔʃ] 'to postpone'
promover [pru]/*[prɔ] 'to promote'

Notice that the contrast between words with the stressed and the unstressed forms of the prefix, shown in (265), correlates positively with the compositional and non-compositional meaning of the relevant words.

- (265) *pré-ocupação* [prɛ] *preocupação* [prj]
 'previous occupation' 'concern'

The behavior of *pre-* when followed by a vowel in this lexicalized use is also distinct from the behavior of the productive stressless prefix with a similar phonological shape *re-*. In fact, as illustrated in (266), in the latter case there are two possible realizations for the prefix: the most common realization is with a [j] (or a [i]), but, in a more careful pronunciation, a schwa may also surface. The latter realization is very odd, arguably impossible, with *pre-*.

- (266) *reordenar* [j]/[i] 'to rearrange'
preocupação [j]/*[i] 'concern'

Interestingly, the behavior described for the incorporated prefix is also found in word internal position, where a glide that corresponds to an underlying non-back vowel may not surface as a schwa, as illustrated in (267).

- (267) (*passEio* [ɐ] < /e/) *passeAr* [j]/*[i] 'to walk'
 (*arEia* [ɐ] < /e/) *areAl* [j]/*[i] 'sand dune'

This shows that *pre-* behaves like prosodic word internal, whereas *re-* does not. The distinction just observed is in fact an important one, since when a

given form may either have a compositional and a non-compositional meaning, like *reunião*, the realization of the prefix with a schwa will always imply the compositional meaning ('reunification'), while the realization with a glide will be ambiguous (meaning either 'reunification' or 'meeting').

To sum up, it can be concluded that the productive stressless prefixes *re-* and *des-* are prosodized as adjuncts to the prosodic word that includes their morphological base, at least when they occur in transparent words. Non-transparent prefixed words, by contrast, may be prosodized similarly to simple words.

4. Host plus clitic combinations

Like derived words with affixes that do not form stress domains independent of their morphological bases, host plus clitic combinations also constitute a domain with a single primary stress.

Besides the issues related to the prosodic structure associated to clitic-host combinations, which are common to derived and inflected words with a single primary stress, as we will see below in section 4.3, there are two additional questions that pertain to clitics proper. These questions concern the (phonological) definition of this class and the lexical treatment of clitic words. These issues are addressed in sections 4.1 and 4.2, respectively.

4.1. Phonological characterization of clitic words

4.1.1. On the definition of clitics

Languages possess different types of *particles* whose classification is difficult to set up. Attempts to establish such a classification include Zwicky and Pullum (1983, 1985), where a number of tests for the identification of clitics, particles, and affixes are proposed. It is frequently noted, nevertheless, that certain elements show properties that can be associated to different classes. Romance pronominal clitics have been pinpointed as an example of this kind, since they show some syntactic mobility, like clitics, but have selection requirements that are typical of affixes (e.g. Klavans 1985; Zwicky 1987a; Spencer 1991: chap.9; Halpern 1995; Monachesi 1996; Miller and Sag 1997).²³⁴

Within the class of clitics, it is also possible to find different subclasses. In the influential study of Zwicky (1977), three categories are proposed: *simple clitics*, *special clitics*, and *bound words*.

Simple clitics correspond to free morphemes that may appear phonologically reduced when they are unaccented, while appearing in the full form in accented positions (e.g. under emphasis or in isolation); the phonological relationship between the full form and the reduced form is in general transparent; the reduced form has the same syntax as the full form and is more frequent in casual style. Typical examples of this type of clitics are the English reduced variants of object personal pronouns.

Special clitics correspond to unaccented variants of a stressed free form; they are weak and prosodically dependent; their phonological form may not be clearly related to the phonological form of the strong counterpart; the selection of either the clitic or the strong form usually follows from syntactic and semantic requirements; and they often show a special syntax. Examples of special clitics are clitic personal pronouns in many Romance and Slavic languages.

Bound words correspond to elements that are always unaccented and dependent on adjacent material, although showing some syntactic freedom and appearing associated with words of different morphosyntactic categories. Examples of this type of words are the Latin conjunction *–que* ('and') or the English possessive marker.

There are three major properties shared by the three classes: the (phonological) clitic is always an unstressed element, which in addition shows a certain degree of syntactic freedom, and it belongs to the class of function words.²³⁵ These properties may thus constitute broad criteria for the clitic status of a given element.

Table 3 summarizes the distinctive properties of the three types of clitics.

Table 3. Distinctive properties of the three types of clitic words proposed in Zwicky (1977).

Property Clitic type	Reduced variants	Unaccented	Special distribution
Simple clitics	yes	no	no
Special clitics	no	yes	yes
Bound words	no	yes	no

Other properties that are often ascribed to clitics, but that either have exceptions or are not sufficient to identify this class are the following:

- Clitics are at most monosyllabic (e.g. Neijt 1985).²³⁶ However, different languages have been reported to show disyllabic (phonological) clitics: for example, according to Inkelas (1990) in Serbo-Croatian there are disyllabic clitics which are distinct from non-clitic words in that they lack both stress and tone; and Van der Leeuw (1997) cites three other languages with these kind of clitics, namely, Ancient Greek, Latin and Dyirbal.
- Clitics are members of closed classes (e.g. Selkirk 1984). This is the case, for example, of all elements in the long list of categories that may include clitic elements presented in Zwicky (1977: 10). Nevertheless, not all units that belong to closed classes are necessarily clitic, and this may vary crosslinguistically.²³⁷
- Clitics cannot appear in isolation. This is true of clitics, but also of other elements that have the properties of prosodic words (for example, of certain prepositions, conjunctions and complementizers in EP, as we will see in the next subsection).

Bearing these properties in mind, in the following section we propose a classification of EP words that have a grammatical function in the language and lack word stress.

4.1.2. Identification and classification of EP clitic words

Our primary aim in this section is to determine which words among the set of EP function items are stressless.

A major cue for the lack of word stress is the process of vowel reduction (see chapter 3, section 3). In particular, the presence of the vowels [i] and [ɐ], which do not belong to the underlying system of EP and which in general do not appear in stressed positions, is a clear symptom of a stressless element.²³⁸ According to this criterion, the monosyllabic function words listed in (268) may be classified as stressless.²³⁹

(268) Monosyllabic words with [i]

de, me, te, se, lhe(s), que, se, que

‘of’, ‘to me’, ‘to you’, ‘to oneself’ ‘to him/her/(them)’, ‘that’, ‘if’, ‘which’

176 *The prosodic word*

Monosyllabic words with [ɐ]

a(s), a, da(s), a(s), na(s), ma(s), ta(s), lha(s), mas

‘the-FEM(PL)’, ‘to’ (preposition or complementizer), ‘of-the-FEM(PL)’, ‘her/it-FEM(PL)’, ‘in-the-FEM(PL)’, ‘to-me-it-FEM(PL)’, ‘to-you-it-FEM(PL)’, ‘to-him/her/them-it-FEM(PL)’, ‘but’

The same criterion implies the clitic status of the disyllabic function words in (269), since the two vowels they are composed of are [ɐ].

- | | | |
|-------------------|--------|----------|
| (269) <i>para</i> | [pəɐɐ] | ‘for/to’ |
| <i>cada</i> | [kədɐ] | ‘each’ |

A phonological process that allows us to demonstrate the stressless status of certain monosyllabic function words ending with a high vowel is V1 semivocalization (see chapter 3, section 11.1). The possibility of V1 semivocalization is illustrated in (270).

- | | | |
|------------------|-----|-------------------|
| (270) <i>que</i> | [j] | ‘that (he) loved’ |
| <i>e</i> | [j] | ‘and animals’ |
| <i>do</i> | [w] | ‘of the author’ |

Besides all the monosyllabic words ending in schwa – which may in general be realized with a glide when followed by a word starting with a vowel²⁴⁰ – the monosyllabic function words that allow semivocalization and that must therefore be unstressed are listed in (271), together with their respective plural forms.

- (271) *o(s), do(s), no(s), o(s), mo(s), to(s), lho(s)*
 ‘the-MASC(PL)’, ‘of-the-MASC(PL)’, ‘in-the-MASC(PL)’, ‘him/it-MASC(PL)’, ‘to-me-it-MASC(PL)’, ‘to-you-it-MASC(PL)’, ‘to-him/her-it-MASC(PL)’

The clitic status of these function words ending with schwa or a round vowel is further corroborated by the fact that such items do not undergo the processes of vowel deletion that operate in prosodic word final position (described in chapter 3, sections 12 and 13). This is illustrated in (272) and (273). The examples in (272a) show the (nearly) obligatory deletion of stressless non-back vowels in prosodic word final position, while the forms in (272b) illustrate the non-application of this process to clitic words. The

examples in (273a) show the optional application of round vowel deletion to vowels in prosodic word final position, which does not apply to clitic words, as illustrated in (273b).

- | | | | | | |
|----------|-------------------------|--------------------|----|---------------------|-------|
| (272) a. | <i>rede aberta</i> | ^x [j]/0 | b. | <i>de abertura</i> | [j]/0 |
| | ‘open net’ | | | ‘of opening’ | |
| | <i>ataque agressivo</i> | ^x [j]/0 | | <i>que agradava</i> | [j]/0 |
| | ‘violent attack’ | | | ‘that pleased’ | |
-
- | | | | | |
|----------|----------------------|-------|--------------------|--------|
| (273) a. | <i>ano anterior</i> | [w]/0 | <i>no anterior</i> | [w]/*0 |
| | ‘previous year’ | | ‘in the previous’ | |
| | <i>medo absoluto</i> | [w]/0 | <i>do absoluto</i> | [w]/*0 |
| | ‘absolute fear’ | | ‘of the absolute’ | |

Notice that the possibility of the clitic words in (272b) to lose their only vowel may be explained by a more general property of (frequent) function words to undergo reduction processes. At least three properties distinguish the process of non-back vowel deletion from this type of reduction: the former is nearly obligatory in unstressed environments, it may be blocked if a following vowel bears word stress (see chapter 6, section 3), and its output is categorical; whereas the latter is optional, not sensitive to the presence of stress in the following vowel, and gradient, as shown in detail in chapter 7.

The fact that the (nearly) obligatory process of final non-back vowel deletion does not apply to the function word *porque* ‘why/because’ (interrogative/relative pronoun, complementizer), as illustrated in (274), indicates furthermore that this disyllabic item also behaves like a clitic word (see further below other arguments in this direction).

- | | | |
|-------|---|-------|
| (274) | <i>Porque andas tão triste?</i> | [j]/0 |
| | ‘Why are (you) so sad?’ | |
| | <i>Luto porque acho importante.</i> | [j]/0 |
| | ‘(I) fight because (I) think it’s important.’ | |

The optional realization of the function word *ou* ‘or’ as [o] or [ɔ] may indicate that this word is also unstressed, since stressed vowels do not allow this type of variation. By contrast, as we have seen in chapter 3, this type of alternation may affect stressless vowels, particularly in word initial position. This certainly points to the unstressed status of this function word.

As already mentioned, certain function words may in addition undergo reduction processes, which result in the loss of segmental material, includ-

ing in some cases their only vowel. Although this is the topic under investigation in chapter 7, we list in (275a) some of the shapes these function words may take (see also chapter 7 for the conditions under which these forms may appear). Such a behavior distinguishes these words from full items, whose stressed vowels are not affected by these reduction phenomena (see 275b).

- | | |
|--|--|
| (275) a. <i>para</i> [pəɾə], [pɾə], [pə]
‘for/to’
<i>em</i> [ẽj], [ẽ]
‘in’
<i>pelo</i> [pelu], [plu], [pl]
‘by-the-MASC’
<i>com</i> [kõ], [ko], [kw], [k]
‘with’
<i>de</i> [dĩ],[dj], [d]
‘of’
<i>ao</i> [aw], [ɔ]
‘to-the-MASC’
<i>por</i> [pur],[pr]
‘by’ | b. <i>pára</i> [páɾə], *[pɾə],*[pə]
‘(he) stops’
<i>mãe</i> [mẽj], *[mẽ]
‘mother’
<i>pêlo</i> [pélu], *[plu],*[pl]
‘hair’
<i>som</i> [só], *[so], *[sw], *[s]
‘sound’
<i>dê</i> [dé],[dĩ],[dj],*[d]
‘give’
<i>pau</i> [páw], *[pɔ]
‘stick’
<i>Artur</i> [ɐrtúr],[ɐtr]
‘id.’ |
|--|--|

That the function words considered so far are not prosodic words is further supported by the fact that they do not have to conform to prosodic word phonotactic restrictions: in particular, the forms of the dative pronoun – *lhe(s)/lho(s)/lha(s)* – start with [ʎ], a segment that cannot appear in prosodic word initial position.

Also pointing in the same direction is the absence of intonational phenomena typical of prosodic words. In fact, judging from the examples of pitch contours of sentences with these function words present in much of the literature on EP intonation (e.g. Viana 1987; Frota 1991, 2000; Falé 1995; Vigário 1998a), such words are not assigned pitch accents.²⁴¹ Notice that, according to Vigário (1998a) and Frota (2000), pitch accent assignment to (phonological) words that are not the head of ϕ is possible in EP. Thus, the absence of pitch accent associated to these elements may follow from their unstressed status, given that pitch accents are only associated to prosodic word stressed positions (see chapter 3, section 18). The same

explanation accounts for the impossibility of assigning focal stress to these words (see chapter 3, section 19).

Another piece of evidence for the distinction between stressed and unstressed (function) words comes from the impossibility of stressless function words to become the prominent element of a phonological phrase by a change in the unmarked stress pattern of ϕ , a possibility available at least to some stressed items, such as the sentence negation marker *não* (cf. Vigário 1997a, 1998a).

We list in (276) the elements that exhibit a clitic behavior in EP, according to the data presented in the previous paragraphs.

(276) List of EP clitic words

Monosyllabic clitics

<i>a, de, por, com, em</i>	prepositions
<i>o(s), a(s)</i>	definite articles
<i>me, te, se, lhe(s), nos, vos, o(s), a(s)</i>	personal pronouns
<i>e, mas, ou</i>	conjunctions
<i>que, se, de, em, por, a</i>	complementizers
<i>que</i>	interrogative/relative pronoun
<i>do(s)/da(s); no(s)/na(s); ao(s)/à(s)</i>	prepositions+definite articles
<i>mo(s)/ma(s); to(s)/ta(s); lho(s)/lha(s)</i>	accusative+dative pronouns

Disyllabic clitics

<i>para</i>	preposition, conjunction
<i>pelo(s)/a(s)</i>	preposition+definite article
<i>cada</i>	quantifier, pronoun
<i>porque</i>	interrogative/relative pronoun, complementizer

It is possible that this list is not exhaustive since there are other elements that may be unstressed but whose phonological behavior cannot be easily differentiated from the behavior of stressed words. A few other forms that may belong to the class of lexically unstressed elements are, for instance, (i) the conjunctions *pois* ‘as’ and *nem* ‘nor’, which do not undergo vowel reduction because these words contain a diphthong and vowel reduction does not apply in such an environment (see chapter 3, section 3), and (ii) the indefinite article *um* [ũ] / *uma* [umə] (‘a-MASC/FEM’).

On the other hand, it is clear that there are also words in EP that belong to some of the morphosyntactic classes listed in (276) but that do not pattern like clitic words. A few examples are provided in (277) where, except for relative/interrogative pronouns which have a diphthong or a nasalized vowel, the remaining forms show stressed vowels that may not have undergone vowel reduction.

(277) Some stressed function words

<i>após, até, desde, sobre</i> 'after', 'until', 'since', 'on'	prepositions
<i>ele(s), ela(s), nós, nosso(s), nossa(s)</i> 'he/them-MASC(PL)', 'she/them-FEM(PL)', 'us', 'our-MASC(PL)', 'our-FEM(PL)'	personal pronouns
<i>logo, embora, caso, conforme</i> 'then', 'although', 'if', 'as'	conjunctions/ complementizers
<i>todos</i> 'all'	quantifier
<i>quem, onde</i> 'who', 'where'	relative/ interrogative pronouns

As for the relative/interrogative pronouns included in the list presented above, we must find other sources of evidence for their stressed status, since, apart from *que*, all relative/interrogative pronouns have as a potentially stressed vowel either a non-central high vowel (which may occur both in stressed and in unstressed positions) or vowels that appear in contexts that do not allow the application of vowel reduction (e.g. vowels in nasal environments or in syllables closed by /l/, and diphthongs – see chapter 3, section 3). In these cases, therefore, vowel reduction may not be used to evaluate the (un)stressed status of the relevant words. Here, the distribution and shape of interrogative pronouns may be revealing, as the following discussion will show.

We have included the interrogative/relative pronoun *que* 'which' in (276) because this word has a schwa as its only vowel. Since this vowel never occurs in stressed position in EP we were led to conclude that such a word must be stressless. The interrogative pronoun *que* has a specificity that sets it apart from both other clitic words in EP and most other interrogative pronouns: it alternates with the strong form *quê* ([ké]),

depending on its position in a sentence. This is illustrated in (278): in (278a) the Wh-word is in sentence initial position and the selected form is necessarily *que* ([ki]), whereas in (278b) it is *in situ* in sentence final position and the selected form is obligatorily *quê* ([ké]).²⁴²

- (278) a. *O que viste tu?* [ki] **O quê viste tu?* [ké]
 b. *Tu viste o quê?* [ké] **Tu viste o que?* [ki]
 ‘What did you see?’

Using the possibility of occupying the sentence final position as a diagnostic for the stressed status of interrogative words, we may establish that besides *que*, only *porque* ‘why’ is unstressed (as the absence of non-back vowel deletion had already indicated), while the remaining Wh-words are stressed. This is illustrated in (279).²⁴³

- (279) a. *Porque fizeste isto?* [purki] *Fizeste isto porque?* [purké]
 ‘Why did (you) do this?’
 b. *Quem encontraste tu?* [kẽj] *Tu encontraste quem?* [kẽj]
 ‘Who did you meet?’
 c. *Onde vais tu?* [õdi] *Tu vais onde?* [õdi]
 ‘Where are you going?’

The same type of classification obtains if we consider the Wh-words that can appear in isolation. Indeed, almost all interrogative words may occur in isolation with the same form as they appear in sentence initial position, to the exception of (o) *que* and *porque*, which must occur in the strong forms *o quê* [uké] and *porquê* [purké].

These are the only cases in EP where stressless and stressed variants of a function word are selected according to their position. Such paucity of weak/strong variants whose occurrence depends on the context appears to be similar to Neapolitan where, according to Bafille (1997: 475), among all clitic function words only the complementizer *que* shows the alternation between [ke] and [ki]. This contrasts with many Germanic languages where this sort of alternations are common (cf. Selkirk 1984: chap.7 and Inkelas 1990: 8.1 for English; Booij 1996a: 223–224 for Dutch; Wiese 1996: 7.4.3 and Hall 1999a: 104 for German; see also chapter 7, sections 5.1 and 5.5 on the phonological relation between reduced and full forms in EP).

To conclude this section, there is substantial evidence for the clitic status of the function words listed in (276). Under Zwicky's (1977) classification reviewed above, these words may fall in the following categories: stressless pronominal clitics have a special distribution (already shown in chapter 4) and thus belong to the class of special clitics; function words with schwa or [ə] as their only vowels are necessarily lexically unstressed, and thus may confidently be classified as bound words; the remaining function words behave like items lacking word stress, and thus may either be lexically unstressed or postlexically reduced – they may, therefore belong either to the class of simple clitics or of bound words. All of these elements may, furthermore, be classified as phonological clitics in the sense that they are stressless elements that need a host to lean on.

For the purposes of the present chapter, the data surveyed in this section plays an important role in that it enables us to establish which function words pattern as clitics in EP, and thus to proceed with our investigation on the prosodization of clitic words.

4.2. *Lexical prosodization of clitics*

It is often considered that clitic function items may be directly inserted in the postlexical component, without undergoing lexical phonological processes (cf. Kaisse and Shaw 1985: 9). Such a view is explicitly defended in Inkelas (1990: 241), who suggests that clitics lack the properties of full words when these properties are assigned by rules that apply at the lexical level, and this is so because clitics do not undergo lexical rules. However, EP clitics undergo phonological rules that have the properties of a lexical process, like vowel reduction and final nasal diphthongization (see chapter 3, sections 3 and 5, respectively). This shows that clitics, like other words, may be affected by processes that are only operative within the lexical component. By the same token, if clitics are present in the lexical component this implies that, like other words, clitics may already be prosodized in this component (see chapter 1, section 3.2 on the construction of prosodic domains up to the prosodic word level within the lexical component). Given this assumption, the present section is devoted to the investigation of the lexical prosodization of clitics.

Rather than positive evidence for the way clitics are prosodized in the lexical level, EP only seems to provide clear evidence for the way clitics *are not* lexically prosodized. In fact, as we have already seen, it is clear that clitics do not have the properties of lexical prosodic words, since they lack

primary word stress and thus their vowels undergo vowel reduction. In addition, they do not have to conform to the phonotactic constraints that hold for lexical prosodic words, since some pronominal clitic forms start with [ʎ], whereas this segment does not occur in prosodic word initial position.

As for the other lexical prosodic domains – the foot and the syllable – little can be said. With respect to the *foot*, little can be said in general, since, as far as we know, there is no clear evidence for this prosodic domain in EP (see the discussion on the visibility of prosodic domains in chapter 8). As for the *syllable*, we can observe that the words classified as clitics in the preceding section show a lexical make up that is consistent with their lexical parsing into syllables. Indeed, the set of syllable structures attested for clitic words is a subset of the set of syllable structures found in the language (cf. d’Andrade and Viana 1994; Vigário and Falé 1994; Mateus and d’Andrade 2000), as illustrated in (280) and (281), below.²⁴⁴

(280) Syllable composition of clitics		Similar composition in syllables of full words		
V	<i>o</i>	V	<i>ó</i>	‘name of the letter <i>o</i> ’
CV	<i>da</i>	CV	<i>dá</i>	‘(he) gives’
CVC	<i>por</i>	CVC	<i>pôr</i>	‘to put’
VC	<i>as</i>	VC	<i>ás</i>	‘ace’
VG	<i>ao</i>	VG	<i>eu</i>	‘I’
(CVGC)	<i>pois</i>	CVGV	<i>dois</i>	‘two’

Only a few types of syllables that are allowed in full words are not attested in clitic words, as those in (281).²⁴⁵ However, these are probably just accidental gaps.²⁴⁶

(281) CCV	<i>crê</i>	‘believe’	CCVG	<i>breu</i>	‘tar’
CCVC	<i>crer</i>	‘to believe’	CCVGC	<i>grãos</i>	‘grains’

We have not found lexical processes that are bound by the syllable domain in EP, which would enable us to be certain that clitics are/are not parsed into syllables within the lexical component. The conclusion we may draw at this point is that clitic words that undergo vowel reduction are not lexically prosodized as prosodic words, but they may be lexically syllabified and thus enter the postlexical level as syllable constituents.

4.3. *Postlexical prosodization of clitics*

Assuming that the prosodic hierarchy does not include the clitic group (cf. Inkelas 1990; Selkirk 1996; Booij 1996a; Peperkamp 1997a; see also the review in chapter 1, section 4), we will explore now the way clitics are postlexically integrated into prosodic trees.

Under this topic, there are several issues to be considered. First of all, it was seen in chapter 1, section 6.3 that clitics may show different directions of cliticization, depending, among other factors, on the general tendency of the language, on the specificity of particular clitics, or on the position of the clitic with respect to adjacent potential hosts (e.g. Booij 1996a; Kleinhenz 1996). With respect to EP, some researchers have suggested that the language has a general tendency for enclisis (cf. Brandão de Carvalho 1989; Galves and Galves 1995; Van der Leeuw 1997; see also chapter 2 for a review). Thus, one of our tasks will be to determine the direction of cliticization of EP clitics in order to evaluate these proposals.

Clitics have also been reported to lean on different prosodic hosts, among which are the prosodic word, the phonological phrase and the intonational phrase (cf. Inkelas 1990; Zec and Inkelas 1991; Selkirk 1996; Booij 1996a; Kleinhenz 1996; Peperkamp 1997a; Hall 1999a, among others – see also chapter 1, section 6.3). The second major issue investigated in this section concerns the possible hosts for cliticization in EP.

Finally, the way clitics are integrated into prosodic structure and the prosodic configuration they display seems also to vary cross-linguistically (e.g. Selkirk 1996; Booij 1996a; Kleinhenz 1996; Peperkamp 1997a; Hall 1999a; see also chapter 1, section 6.3). Thus, at the end of this section we hope to be able to add the EP case to the more global picture of the available possibilities for the prosodization of clitics.

To accomplish these goals, we will systematically use the phonological phenomena classified as postlexical in chapter 3, which we list in (282) below for ease of reference (see also Table 2 of section 2). This procedure follows from our previous observations: as stressed in section 2 above, since we argue that clitics are only combined with their hosts postlexically (see chapter 4), except for precompiled phrasal phenomena, the processes that apply within the lexical component are predicted never to be affected by the postlexical combination of host and clitic (see chapter 4, section 3.2 for the demonstration that indeed they never are).

- (282) Postlexical phonological phenomena that refer to the prosodic word domain

- Initial vowel realization (feature specification)
- V2 semivocalization
- Final non-back vowel deletion
- Final round vowel deletion
- Final central vowel deletion
- Syllable degemination
- Initial stress
- Emphatic stress
- I-initial H-tone assignment

Among such phenomena, those that refer to the left-edge of the prosodic word (listed in 283a) will be crucial to determine which clitics (if any) are proclitic to the prosodic word domain and under which prosodic configuration;²⁴⁷ conversely, those that refer to the right-edge of the prosodic word domain (listed in 283b) will be used as diagnostics for the identification of the EP clitics (if any) that lean on the preceding prosodic word, as well as the type of prosodic structure they display.

- (283) a. Processes that refer to prosodic word initial position

- Initial vowel realization (feature specification)
- Initial stress
- Emphatic stress

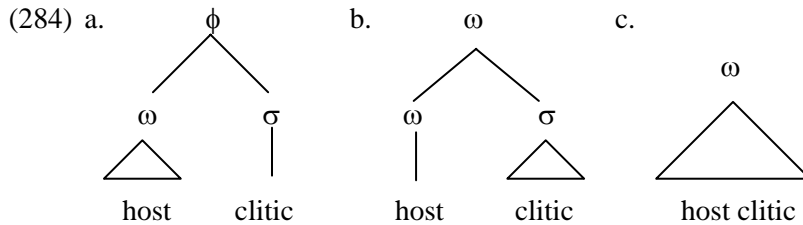
- b. Processes that refer to prosodic word final position

- V2 semivocalization
- Final non-back vowel deletion
- Final round vowel deletion
- Final central vowel deletion
- Syllable degemination

The following two sections are devoted to the investigation of the prosodization of enclitics (section 4.3.1) and of proclitics (section 4.3.2).

4.3.1. Prosodization of enclitics

According to several proposals reviewed in chapter 1 (e.g. Selkirk 1996; Booij 1996a; Kleinhenz 1996; Peperkamp 1997a; Hall 1999a), enclitics may be assigned at least one of the configurations in (284).



We will try to demonstrate that in EP only postverbal pronominal clitics behave like phonological enclitics. We further hope to show that enclitics behave like units that are incorporated into the prosodic word that dominates their host, as in the configuration (284c).

Let us consider V2 semivocalization, a process described in chapter 3, section 11.1 to apply optionally to stressless round high vowels in prosodic word final position when preceded by a vowel, as in (285a). The examples in (285b–c) illustrate that V2 semivocalization may apply to postverbal pronominal clitics (285b), but not to preverbal pronominal clitics or to the definite article (285c).²⁴⁸

(285) a. <i>o rio lento</i>	[u]/[w]	‘the slow river’
b. <i>eu vi-o deitado</i>	[u]/[w]	‘I saw him laid down’
<i>eu li-os todos</i>	[u]/[w]	‘I read them all’
c. <i>só a ti o diz</i>	[u] ^x /[w]	‘(he) only tells it to you’
<i>eu vi o texto</i>	[u] ^x /[w]	‘I saw the text’

This constitutes a first indication that postverbal pronominal clitics, but not preverbal pronominal clitics or the definite article, are hosted by the prosodic word that includes the verb.

A clearer test for the prosodization of clitics is final non-back vowel deletion, since the intuitions regarding this process are very strong and the (non-)application of the process is very neatly perceived. This rule consists of the deletion of a non-back stressless vowel in prosodic word final posi-

tion, as illustrated in (286) (see also chapter 3, section 12 and the references therein).

- | | | | |
|-------|-------------------|---------------------|--------------|
| (286) | <i>pele alva</i> | 0/ ^x [j] | ‘white skin’ |
| | <i>bebe agora</i> | 0/ ^x [j] | ‘drink now’ |

As the data in (287a) shows, if a pronominal clitic appears in postverbal position, the final vowel of the verb is not deleted. By contrast, the final non-back vowel of the postverbal clitic in (287b) must be deleted. Recall that in non-postverbal position clitic words ending with a non-back vowel show the alternation [j]/0. That is, as we have already seen, non-back vowel deletion – which is a nearly obligatory process – does not apply to clitic function words.

- | | | | |
|----------|-----------------------|---------------------|--------------------------|
| (287) a. | <i>pede-o já</i> | *0/[j] | ‘ask it now’ |
| | <i>mede-a depois</i> | *0/[j] | ‘measure it later’ |
| b. | <i>peço-te agora</i> | 0/ ^x [j] | ‘(I) ask you now’ |
| | <i>disse-me ontem</i> | 0/ ^x [j] | ‘(he) told me yesterday’ |

These facts demonstrate that, when the verb is followed by a pronominal clitic, its final vowel does not behave as prosodic word final (since in this case vowel deletion becomes impossible); conversely, the pronoun behaves as a prosodic word final element, although it does not constitute a prosodic word independent of its host. Of the three possibilities of prosodization presented above, only (284c), which involves the prosodic incorporation of the clitic into the verb’s prosodic word, is compatible with these facts. The prosodization of the forms in (287) is thus proposed to be the one in (288).

- | | | |
|-------|-----------------------|-------------------------|
| (288) | (pede-o) _ω | (peço-te) _ω |
| | (mede-a) _ω | (disse-me) _ω |

Let us now consider preverbal pronominal clitics and other clitic words ending in a schwa or starting with a vowel in order to determine whether the same prosodization may be proposed in such cases.

The forms in (289a) show that non-back vowel deletion applies to the full word that precedes other clitics. This means that the full word corresponds to a prosodic word, as represented in the partial structure provided in (289b).

- (289) a. *pede o livro* 0/^x[j] b. (pede)_ω o livro
 ‘ask the book’
 só Henrique o disse 0/^x[j] só (Henrique)_ω o disse
 ‘only Henrique said it’

Putting it differently, these clitic words are not incorporated into the preceding prosodic word, unlike postverbal clitic pronouns. Finally, the examples in (290a) suggest that these clitic words are not adjunct to the prosodic word domain either, since otherwise they should undergo the process of non-back vowel deletion, contrary to fact. The ruled-out structure is provided in (290b).

- (290) a. *apenas te ofereci* 0/[j] b. *((apenas)_ω te)_ω ofereci
 ‘(I) only offered you’
 gosto de observar 0/[j] *((gosto)_ω de)_ω observar
 ‘(I) like to watch’

The alternation 0/[j] in the realization of clitics other than postverbal pronouns could be thought, instead, to be a consequence of the bidirectionality of these clitics – as proposed for languages like Dutch or German (cf. Booij 1996a and Kleinhenz 1996, respectively). We have to reject this possibility, however, because the same alternation is also found in intonational phrase initial position, where no host to the left is available (see chapter 7 for our account of this alternation).

The data considered so far converge with respect to the direction of cliticization of clitic words, as well as to the host they lean on. Postverbal pronominal clitics pattern phonologically like enclitics to the preceding prosodic word, while the remaining clitic words do not exhibit such behavior. In addition, non-back vowel deletion strongly supports the prosodic incorporation of pronominal enclitics into the prosodic word that includes the verbal host.

The same global picture obtains as far as the process of final round vowel deletion is concerned. As illustrated in (291a–a’), this process optionally applies to stressless round vowels in prosodic word final position (see chapter 3, section 13). The example in (291b) shows that *postverbal* pronominal clitics behave like prosodic word final elements (although they do not constitute independent prosodic words). The partial prosodic structure that is compatible with this behavior is shown in (291b’). Finally, the sequence in (291c) confirms the prosodic incorporation of the clitic into

the preceding prosodic word, since the verb's final vowel may no longer be deleted in this configuration. The prosodization of the relevant sequence is shown in (291c').

- (291) a. *um belo amigo* [w]/0 a'. um (belo)_ω amigo
 'a nice friend'
 b. *posso vê-lo agora* [w]/0 b'. posso (vê-lo)_ω agora
 '(I) can see him now'
 c. *embalo-a já* [w]/*0 c'. (embalo-a)_ω já
 '(I) rock her now'

The application of this process to the prosodic word that precedes other clitics, as well as the impossibility of other clitics to undergo round vowel deletion corroborates our preceding observations pointing to the non-enclitic nature of these clitic elements, as illustrated in (292). In (292a) the optional rule may apply to the final vowel of the full word preceding the clitic, thus indicating that the clitic does not incorporate into the preceding prosodic word. This is shown in the partial prosodic structure in (292a'). In (292b) the process may not apply to the clitic word, which constitutes a proof that this element is not adjunct to the preceding prosodic word either. The ruled-out prosodic configuration is given in (292b').

- (292) a. *embalo a bebé* [w]/0 a'. (embalo)_ω a bebé
 '(I) rock the child'
 b. *gosto do artigo* [w]/*0 b'. *((gosto)_ω do)_ω artigo
 '(I) like the article'

Central vowel deletion also shows the same pattern. This process applies to stressless central vowels in prosodic word final position when followed by a word starting with a non-high non-central vowel, as illustrated in (293a) (see chapter 3, section 14). A postverbal pronominal clitic may also undergo this process, as shown in (293b). This implies that the clitic is attached to the preceding prosodic word (293b'), since clitics do not form independent prosodic words and thus may not independently undergo rules that apply to the prosodic word domain.²⁴⁹

- (293) a. *ele pensa objectivamente* [ɐ]/0
 a'. ele (pensa)_ω objectivamente
 'he thinks objectively'

- b. *a gata, pude vê-la outrora* [v̥]/0
 b'. *pude (vê-la)_ω outrora*
 'as for the cat, (I) could see her in the past'

Final central vowel deletion provides further evidence for the non-enclitic status of the remaining clitic words, including preverbal pronominal clitics, since it may not apply to the clitic itself (see 294a–b). This entails that such clitics are not attached to the preceding prosodic word, as indicated in the partial prosodic structures in (294a'–b').

- (294) a. *falo da organização* [v̥]/*0 a'. **(falo da)_ω organização*
 '(I) speak of the organization' *((falo)_ω da)_ω organização
 b. *apenas ta ofereci* [v̥]/*0 b'. **(apenas ta)_ω ofereci*
 '(I) just gave it to you' *((apenas)_ω ta)_ω ofereci

We should notice, nevertheless, that there is another phenomenon that interacts with central vowel deletion which somewhat obscures the facts in this case. The monophthongization/merging of stressless [aw]/[v̥u] yielding [ɔ] often occurs with a verb and an enclitic (295a), a word and a proclitic (295b), two clitics (295c), the clitic word *ao* (295d), and within some full words (295e) (some of these possibilities are already described in Gonçalves Viana 1883: 37).

- (295) a. *diga-o você* [v̥u]/[ɔ] 'say it yourself'
 b. *diga o necessário* [v̥u]/[ɔ] 'say the necessary'
 c. *para o jantar* [p̄v̄v̄u]/[prɔ]/[pɔ] 'for (the) diner'
 d. *diz ao João* [aw]/[ɔ] 'tell (to-the) John'
ao João, já o disse [aw]/[ɔ]
 '(as) to John, I've already told it (to him)'
 e. *automóvel* [aw]/[ɔ] 'car'
restaurante [aw]/[ɔ] 'restaurant'

This process cannot be considered a span rule because the relevant sequence may not merge if the second vowel belongs to a full word, as illustrated in (296) (see also Gonçalves Viana 1883: 37). It seems, furthermore, that it is not a completely general process, as there are some forms where it is very marked or impossible, as in (297).

- (296) *casa usada* [ɐu]/*[ɔ] ‘used house’
bela utopia [ɐu]/*[ɔ] ‘beautiful chimera’
- (297) *autarquia* [ɐu]/^χ[ɔ] ‘autarchy’
auditório [ɐu]/^χ[ɔ] ‘auditorium’
audição [ɐu]/^χ[ɔ] ‘hearing’

Since its application does not allow to distinguish enclitics from proclitics, as the remaining processes do, and given that its domain appears not to be defined on pure prosodic terms (e.g. it applies within the prosodic word, as in the case of the verb plus enclitic, but is also applies to the clitic word *ao*, which does not form an independent prosodic word; and it does not apply uniformly), we believe it may not be used either to support or to contradict any proposal regarding the direction of cliticization.²⁵⁰ Consequently, we will not be concerned with this phenomenon any further.²⁵¹

Syllable degemination also confirms our previous observations. This process, described in detail in chapter 3, section 15, consists of the optional deletion of a vowel and a consonant belonging to the first of two similar syllables, when the first one is unstressed and prosodic word final. Its application when the two words involved are full prosodic words is illustrated in (298a) (from Frota 2000). The impossibility of deleting the final syllable of a verb when followed by a pronominal clitic, as in (298b), is consistent with the analysis that the clitic incorporates into the verb's prosodic word (298b'). This process may also apply to the word that precedes other clitics, as in (298c), thus pointing to the non-incorporation of these clitics into the prosodic word to their left (298c').

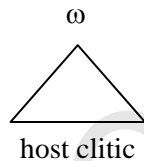
- (298) a. *o campo poluído* [pupu]/[pu] a'. ... (campo)_ω poluído
 ‘the polluted field’
- b. *o João bate-te muito* [titi]/^χ[ti] b'. ... (bate-te)_ω muito
 ‘(the) John hits you a lot’
o livro, meto-to lá [tutu]/^χ[tu] ... (meto-to)_ω lá
 ‘the book, (I) put it (for you) there’
- c. *ele entende de tudo* [didi]/[di] c'. ... *(entende de)_ω tudo
 ‘he knows about everything’
o agrado do João [dudu]/[du] ... *(agrado do)_ω João
 ‘the pleasure of John’

Syllable degemination may not apply when the first of the two adjacent syllables belongs to a clitic word, whether enclitic or not. This may be due to the necessity of preserving the grammatical information conveyed by these function words. Indeed, if the process consists of the loss of the *first* of two adjacent syllables (as assumed, for example, in Sá Nogueira 1938 and Frota 2000 – see also chapter 3, section 15 for evidence in favor of this analysis), the sequences involving a monosyllabic clitic in the first position imply the deletion of the whole function word. Despite this, we should note that syllable degemination is also impossible with disyllabic clitic words, as illustrated in (299). In such cases, we can attribute this impossibility to the prosodic configuration that these clitics display: specifically, to the fact that these units are not independent prosodic words, on the one hand, and are not enclitic to the preceding prosodic word, on the other hand.²⁵²

- (299) *disse isto porque queria* [kiki]/^x[ki]
 ‘(I) said this because (I) wanted (to)’
troquei pelo lugar dele [lulu]/^x[lu]
 ‘(I) changed (my place) for his place’

We can conclude that the processes reviewed in this section clearly indicate that postverbal pronominal clitics are incorporated into the prosodic word that includes the preceding verb.²⁵³ This is depicted in (300), below.

- (300) Postverbal pronominal clitics: incorporation into the preceding ω



As for all remaining clitic words, we found no evidence for their enclitic status. On the contrary, the facts demonstrate that, against the suggestions in Brandão de Carvalho (1989), Galves and Galves (1995) and Van der Leeuw (1997), these elements are *not* attached to the preceding prosodic word (either through incorporation or adjunction to the prosodic word node). As far as we can tell, there is no evidence either for these elements to attach directly to a preceding ϕ . Indeed, we will show in the next section that the phonological behavior of these clitics points to their proclitic nature. Already indicative of their proclitic status is the possibility of clitics to

appear in intonational phrase initial position, that is with no host to their left, as illustrated in (301).

- (301) [Penso que]_I [se tudo correr bem]_I [me vão propor umas férias]_I
 ‘(I) think that, if everything goes well, (they) will give me vacation.’

To conclude our demonstration that, besides postverbal pronominal clitics, EP clitics are not enclitic, we add a final argument that results from the observation of the behavior of these function words when no host is available to their right. Such a configuration necessarily obtains when an intonational phrase boundary follows the function word. In EP these cases may arise because intonational boundaries can be inserted either before or after a complementizer or a relative pronoun, as shown in (302).²⁵⁴

- (302) *Penso que, até onde posso prever, acabo hoje esta tarefa.*
 ‘(I) think that, as far as I can anticipate, (I) finish this task today.’
- a. [Penso]_I [que até onde posso prever]_I [acabo hoje esta tarefa]_I
 b. [Penso que]_I [até onde posso prever]_I [acabo hoje esta tarefa]_I

The relevant configuration for the present discussion is the one in (302b). Here, the presence of intonational phrase prominence and the assignment of the tonal nucleus to the complementizer indicate that it is not incorporated into the preceding prosodic word. In fact, incorporated enclitics are never assigned (intonational phrase) prominence or pitch accent.²⁵⁵

Two possible prosodizations could, nevertheless, account for the behavior of the complementizer in intonational phrase final position: on the one hand, this clitic may be adjoined to the preceding prosodic word, as in (303a); on the other hand, occurring in a (postlexically) stressed position, the clitic may be prosodized as a prosodic word, as in (303b).

- (303) a. ((penso)_ω que)_ω b. (penso)_ω (que)_ω

Our proposal is that only the prosodization illustrated in (303b) is adequate in EP. Although the evidence to choose between the two structures is rather subtle, there are two arguments that rule out the analysis where the complementizer is adjunct to the preceding prosodic word.

In the languages where clitics have been argued to be adjoined to the preceding prosodic word, like Neapolitan (cf. Peperkamp 1997a), a new stress only occurs if the adjunction of clitics adds two syllables, which are

- According to the data in Nespor (1999b), the same type of asymmetry between monosyllabic and disyllabic sequences of clitics exists in languages like Spanish, Greek, or Turkish, whether the stress on the clitic is the primary or the secondary stress of the host plus clitic sequence. This behavior is not shared by EP (clitic) function words in intonational phrase final position (see 305, where ‘T*’ stands for pitch accent). As shown in (305a–b), both monosyllabic and disyllabic function words are assigned I-level prominence and the nuclear pitch accent. In addition, at least with the disyllabic clitic *porque*, the stress may either occur in the first or in the second syllable of the word, as shown in (305c) (where the syllable bearing stress appears in *italic*).

- Besides this, these function words may also (optionally) be the only members of an intonational phrase. This is illustrated in (306), where complementizers form their own I-phrase.²⁵⁷

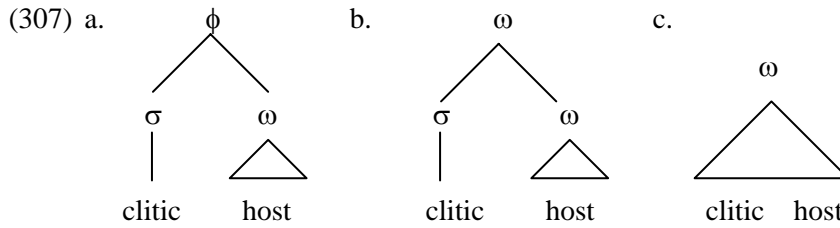
- (306) [Os rapazes louros]_I [que]_I [para a Ana]_I [são os mais giros]_I
 the boys blond which for the Ana are the most cute
 [acabaram de sair]_I
 finished of leaving
 ‘The blond boys, which – for Ana – are the cutest, have just left.’
- [Sei tudo isto]_I [porque]_I [como debes calcular]_I [alguém mo disse]_I
 (I) know all this because as (you) should guess someone to-me told
 ‘I know all of this because, as you may guess, someone told it to me.’

In such cases too, the function words have the same phonological properties as in other intonational phrase final positions: they are the head of the I-phrase, to which the nuclear pitch accent is associated. Here, however, only an analysis where these words form their own prosodic words is available, since there is no other prosodic word that could host them within their intonational phrase. We therefore conclude that, under specific conditions, lexically unstressed words may form independent prosodic words, in particular, when they occur in a stressed position, like the intonational phrase final position.²⁵⁸ This analysis implies, furthermore, that besides postverbal pronominal clitics, EP stressless words are not enclitic. In other words, when function words are clitic in this language they must be proclitic. This will be independently demonstrated in the following section.²⁵⁹

4.3.2 Prosodization of proclitics

All the evidence offered in the previous section for the prosodization of clitics other than postverbal pronouns has shown that they do not attach to the previous prosodic word, thus pointing to the proclitic status of these words. It remains to be seen, however, what kind of prosodic relation the clitic establishes with its host, as well as to identify the prosodic host it leans on.

According to the proposals found in the literature and reviewed in chapter 1, section 6.3, proclitics may be integrated in prosodic structure in at least one of the following ways (cf. Selkirk 1996; Booij 1996a; Kleinhenz 1996; Peperkamp 1997a; Hall 1999a, among others):



In order to establish which of these configurations is selected in EP, we will use as diagnostics for the prosodization of proclitics the processes identified in chapter 3 to refer to the left-edge of the prosodic word: the alternation in height of non-central non-high vowels, prosodic word initial stress assignment and emphatic stress assignment (see chapter 3, sections 10, 16 and 17, respectively).

Before proceeding, however, we would like to make a brief note on other processes that have also been proposed to refer to prosodic word initial position, namely, final round vowel deletion, final central vowel deletion and syllable degemination (see chapter 3, sections 13, 14 and 15, respectively).

As illustrated in (308), the segments belonging to clitic function words behave like prosodic word initial segments as far as the application of these processes is concerned. That is, the processes apply whether the vowels/syllables affected are followed by segments that belong to clitics or to independent prosodic words.

- (308) a. *não visto a camisola* [u]/0
 ‘(I) do not put on the shirt’
 b. *faça chuva ou faça sol* [ɐ]/0
 ‘(whether) it is rainy or it is sunny’
 c. *coitado do João* [dudu]/[du]
 ‘poor (of-the) John’

Nevertheless, as we have already seen in section 3.2 above for syllable degemination, final round vowel deletion and final central vowel deletion may be formalized without reference to prosodic initial position, as in (309) and (310).

- (309) Final round vowel deletion (alternative to the formulation in chapter 3, section 13)

$$\left(\begin{array}{c} \text{V} \\ +\text{round} \end{array} \right) \rightarrow \emptyset \quad / \quad [\dots[\dots_]\omega \text{ V } \dots]_{\text{I}}^{\text{max}}$$

- (310) Final central vowel deletion (alternative to the formulation in chapter 3, section 14)

$$\left(\begin{array}{c} \text{V} \\ -\text{round} \\ +\text{back} \end{array} \right) \rightarrow \emptyset \quad / \quad [\dots[\dots_]\omega \left(\begin{array}{c} \text{V} \\ -\text{high} \\ \alpha\text{back} \\ \alpha\text{round} \end{array} \right) \dots]_{\text{I}}$$

Thus, while the behavior described above is entirely compatible with the prosodization of these elements as proclitic to the right-adjacent prosodic word, it does not actually exclude the possibility that these elements attach to a higher-level prosodic node.

On the contrary, the processes mentioned at the beginning of this section clearly identify the left-edge of the prosodic word in EP, and thus we will concentrate on those processes in the following paragraphs.

The presence of a proclitic word does not seem to affect the realization of prosodic word initial non-central non-high vowels. This is shown in (311), where the realization of word initial vowels varies, whether or not the word is preceded by a proclitic.

- | | |
|-------------------------------------|---------------------------------|
| (311) a. <i>organização</i> [o]/[ɔ] | b. <i>a organização</i> [o]/[ɔ] |
| ‘(the) organization’ | |
| <i>operários</i> [o]/[ɔ] | <i>de operários</i> [o]/[ɔ] |
| ‘(of) workers’ | |
| <i>oceanários</i> [o]/[ɔ] | <i>dos oceanários</i> [o]/[ɔ] |
| ‘(of-the-PL) oceanariums’ | |
| <i>emissão</i> [e]/[i] | <i>a emissão</i> [e]/[i] |
| ‘(the) broadcast’ | |
| <i>elevado</i> [e]/[i] | <i>de elevado</i> [e]/[i] |
| ‘(of) high’ | |
| <i>educador</i> [e]/[i] | <i>o educador</i> [e]/[i] |
| ‘(the) instructor’ | |

This constitutes a first indication that proclitics are not incorporated into the following prosodic word (as in the structure in 307c above), since the initial vowels of the full word that follows the clitic still pattern like (other) prosodic word initial vowels.

Let us now consider the assignment of initial stress. Here, we will use the data presented in Frota and Vigário (2000) (see also chapter 3, section 16). In this pilot experiment 20 sentences were read by two speakers. The experiment consisted of the identification of the perceived prominent syllables of the 40 sentences obtained by each of the authors (excluding the primary word stresses). Considering only the results where both listeners agreed in their responses, which constitute the large majority of cases, we may conclude that the initial stress may either be assigned to the first syllable of a full word or to a preceding proclitic. Examples of the two types are presented in (312), where the prominent (non-lexically stressed) syllables appear in boldface.

- (312) *a autoridade **do** governador **diminuiu***
 ‘the authority of the governor decreased’
*o **organizador** **apresentou** a **catalogadora***
 ‘the organizer introduced the archivist’
*o governador **aceitou** **a** **modernização***
 ‘the governor accepted the modernization’

These facts are strongly suggestive of the prosodization of proclitics as adjuncts to the following prosodic word, as in the representation in (307b) above. Indeed, in such a prosodic configuration, and given that the proclitic is prosodic word initial, the clitic may be assigned the initial stress, as in (313a). Yet, since the syllable starting the host is also in prosodic word initial position, this syllable may also bear initial stress, as in (313b).

- (313) a. (**do** (governador)_ω)_ω
 b. (a (**catalogadora**)_ω)_ω

The fact that in this corpus the same sentence may show the two possible locations of initial stress (as the pair in 314) further supports the proposed analysis.²⁶⁰

- (314) a. *a **inteligência** da **catalogadora** foi determinante*
 b. ***a** **inteligência** **da** **catalogadora** foi **determinante***
 ‘the intelligence of the archivist was crucial’

The same type of alternation in the assignment of initial stress to the beginning of a full word or to a preceding grammatical word is also reported in d'Andrade and Viana (1999). Their analysis, however, differs from ours. For these authors, initial stress is optionally assigned to the first syllable of an *intonational unit* (a constituent similar to the phonological phrase, according to Viana 1987: note 3.3), and this is why the grammatical word may bear a secondary prominence. By contrast, the secondary prominence on the first syllable of the full word that follows the clitic results from the alternations of the rhythmic wave (echo stresses) that are computed from the primary stress leftwards. As for the words that show an odd number of syllables to the left of primary stress, the two principles interact: if the optional rule of initial prominence assigns stress to the first syllable of the word, the rhythmic stress that is predicted on the second syllable does not surface because the result would constitute a stress clash; if initial prominence does not apply, the second syllable may bear a rhythmic stress. The two possibilities are illustrated in (315) (based on d'Andrade and Viana 1999: 92): in (315a) the optional initial rule does not apply, and thus only the echo stress surfaces; whereas in (315b) the initial rule applies, and thus the clashing sequence that would result leads to the demotion of the stress on the second syllable of the word.

(315) a. Echo stress b. Initial prominence

*
.
* * . * *
* * * * *

Constantinopla

*
* . . * .
* * * * *

Constantinopla

However, if we understand this proposal correctly, since initial stress is assigned to the initial syllable of an *intonational unit* (a constituent similar to the ϕ -phrase) and given that unstressed grammatical words appear to begin such a constituent, in a sequence where a full word is preceded by a clitic word the analysis illustrated in (315b) is no longer available. Crucially, in this case the beginning of the intonational unit does not coincide with the beginning of the full word. However, both in our data (see 316a) and in the data presented by d'Andrade and Viana (1999) (see 316b) there are full words that show an odd number of syllables to the left of the main stress which are assigned initial stress although preceded by a stressless grammatical word.

- (316) a. *a **inteligÊ**Ncia* ‘the intelligence’
 *discutiu a **gramaticalidAde*** ‘discussed the grammaticality’
 b. *em **ConstantinOpla*** ‘in Constantinople’

Under our account, these facts are straightforwardly explained given the proposal that initial stress is assigned to the first syllable of the prosodic word and that proclitics are adjuncts to the following prosodic word (317a). By contrast, under d’Andrade and Viana’s approach, initial stress – and hence the configuration that leads to stress clash resolution that is responsible for the stress in the first syllable of the full word – is assigned only to the initial syllable of the intonational unit, which corresponds to the clitic, and not to the initial syllable of the full word (317b).

- (317) a. Our approach:

Initial stress is assigned to the initial syllable of a prosodic word

Predictions: (a (**gramaticalidAde**)_ω)_ω

or (a (**gramaticalidAde**)_ω)_ω

- b. D’Andrade and Viana’s (1999) approach:

Initial stress is assigned to the initial syllable of the *intonational unit* (IU)

Predictions: (**em ConstantinOpla**)_{IU}

but *(em **ConstantinOpla**)_{IU}

Besides this, our approach is more compatible with the observed data as it clearly separates (all) initial prominences from echo stresses. In our corpus, initial stress occurs regardless of the number of syllables to the left of main stress, and regardless of the presence of proclitics. In addition, no internal prominences were obtained. The latter observation is corroborated by d’Andrade and Viana’s (1999: 88) suggestion that non-initial secondary stresses are only likely to *emerge in careful speech at the level of the citation form*. If in normal speech non-initial echo stresses do not surface, and if the initial stress is independently required, we can establish the generalization in (318) for EP.

- (318) In regular speech, non-primary prominences result (only) from the assignment of initial stress.

Although we do not have empirical evidence that postverbal pronominal clitics are never assigned initial stress – which is our prediction, given the analysis of these elements as enclitic to the preceding prosodic word – Frota (1994) and d’Andrade and Viana (1999: 92), among others, observe that in EP these elements do not show the rhythmic alternations that can be found, for example, in some Romance languages or dialects, whether one or more enclitics follow the verbal host.²⁶¹ Thus, this is consistent with the account here proposed.

Our final argument supporting the prosodization of proclitics as adjuncts to the following prosodic word comes from the distribution of the emphatic stress. Although this type of prominence is very similar to initial stress, there are two main reasons for us to include it in the present discussion: first, this stress has the pragmatic effect of highlighting the relevant word, and thus is distinct from initial stress; second, it is also clearly distinct from rhythmic stress. Although we believe the possibilities of assigning emphatic stress can be evaluated through introspection, the following description is based on collected data already referred to in section 3.2 above and chapter 3, section 17, and presented in Appendix I, unless stated otherwise.

Like initial stress, emphatic stress may either appear on the first syllable of a full word, as in the examples in (319a), or in the first syllable of a proclitic word, as in the examples in (319b) (where the syllable with emphatic stress is signaled with capital letters in underlying).

- | | |
|--|--|
| (319) a. <i>a</i> <u>CON</u> taminaçãO
‘the contamination’
<i>expressõEs</i> <u>PRE</u> posicionAls
‘prepositional expressions’
<i>serIam</i> <u>NA</u> cionalizAdos
‘(they) would be nationalized’
<i>deficiÊNcia</i> na <u>CON</u> tabilizaçãO
‘discrepancy in the calculation’ | b. <i>construçãO</i> <u>DE</u> <i>representaçõEs</i>
‘construction of representations’
<i>temporAls</i> <u>E</u> <i>não-temporAls</i>
‘temporal and non-temporal’
<u>PA</u> ra <i>situaçõEs</i> <i>atÉlicas</i>
‘for athelic situations’
<i>portANto</i> <u>NA</u> <i>posiçãO</i> <i>localizadOra</i>
‘thus in the locating position’ |
|--|--|

These possibilities are straightforwardly accounted for under the proposal put forward in chapter 3, section 17 that emphatic stress is assigned to the initial syllable of the prosodic word and given the assumption that proclitics are adjoined to the following prosodic word (see 320).

- (320) Emphatic stress is assigned to the initial syllable of a prosodic word
 Borne out predictions: (na (CONtabilizaçãO)_ω)_ω
 or (NA (posiçãO)_ω)_ω

It should be observed that this stress may also appear immediately before the word main stress, as shown in (321), a possibility that rules out any attempt to derive this type of stress from a rhythmically assigned stress.

- (321) *É FREquENte* ‘(it) is frequent’ *estÁ INtActo* ‘(it) is untouched’
 o MELhOr ‘the best’ *em DIrEcto* ‘alive’

Notice, in addition, that this type of prominence may also appear in ϕ -internal position, as shown in (322), where the relevant prosodic structure is also given (see Frota 2000 and chapter 1, section 2 on the algorithms for the construction of the ϕ -phrase in EP).

- (322) ((do (intervAlo)_ω)_ω (DEnotAdo)_ω)_φ pela...
 ‘of the interval denoted by...’
 ((expressÕEs)_ω (DEnotadOras)_ω)_φ de intervalos ...
 ‘expressions denoting intervals...’

Thus, the data undoubtedly exclude the possibility that emphatic stress is assigned to the initial syllable of a constituent higher than the prosodic word.

Furthermore, the following facts argue against the possibility that proclitic words are promoted to a prosodic word under emphatic stress.

Disyllabic clitics under emphatic stress still behave like clitic, and thus unlike prosodic words, as far as the processes of prosodic word final deletion are concerned. This is illustrated in (323) (data from introspection): in (323a), the function word does not undergo non-back vowel deletion, which applies (nearly) obligatorily when the relevant vowel occurs in prosodic word final position; in (323b), syllable degemination cannot apply, whereas in similar conditions the rule applies optionally if the first syllable belongs to a prosodic word.

- (323) a. *Penso fazê-lo PORque acho correcto.* [j]/0
 ‘(I) think I’ll do it because I feel it is right.’
 b. *Troquei PElo lugar do João.* [lulu]/*[lu]
 ‘I changed my place for John’s.’

Additionally, as shown in (324), the reduced variants of clitic words may also be assigned emphatic stress, although they are not likely to form independent prosodic words (data from introspection).

- (324) *Troquei peLU lugar do Pedro.* [plu]
 O João vai AQ concerto. [ɔ]
 ‘(The) John goes to-the concert.’
Levo isto paRA Cascais. [præ]/[pæ]
 ‘(I’ll) bring this to Cascais.’

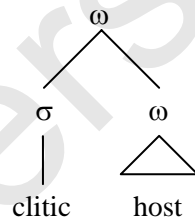
Finally, we should recall that even under emphatic stress the clitic words headed by schwa or [ɐ] are still produced with such vowels and thus the promotion of these function words to prosodic words would (at least) constitute a very marked option.

Despite the fact that more has to be known on these issues before we can use them as clear-cut evidence for the prosodization of function words in EP, we believe they strongly suggest that lexically unstressed function words do not form independent prosodic words under emphasis.

To close this description of emphatic stress distribution, we found in the data no instances of emphatic stress assigned to a word final syllable, and only very rarely did we find non-initial syllables bearing this type of stress.²⁶² According to our intuitions, furthermore, this type of stress may not be assigned to postverbal pronominal clitics. This is consistent with the analysis that these elements are enclitic to the preceding prosodic word, and adds to the sound distinction between postverbal pronominal clitics and the remaining clitic words in EP.

In conclusion, the phenomena observed in this section suggests that, besides postverbal pronominal clitics, the remaining clitic words are adjoined to the following prosodic word, in a structure like (325). The only exception to this is the intonational phrase final position, where these units must form a prosodic word on their own because in this configuration there is no phonological host to their right (see section 4.3.1).

- (325) EP clitics (other than postverbal pronouns):
 adjunction to the following ω



5. Affixes versus clitics

Our investigation on the prosodization of derived and inflected words and of host plus clitic sequences has yielded similar results. On the one hand, in the same way suffixes are included in the same prosodic word as their morphological base, enclitics are incorporated into the prosodic word that dominates the verbal host. On the other hand, prefixes are adjoined to the prosodic word that dominates their morphological base, in the same fashion proclitics are adjoined to the following prosodic word. Nevertheless, we have assumed throughout this book that clitics are distinct from affixes in that the latter are combined with the prosodic word that dominates their morphological base at the lexical level, whereas the former are combined with their prosodic hosts postlexically (see chapter 4 for the demonstration of the latter claim). Some difference between these elements can therefore be expected. In the following paragraphs, we illustrate in a systematic way the distinction between suffixes and enclitics, on the one hand, and prefixes and proclitics, on the other. It will be suggested that this distinction results either from the locus in the grammar where phonological and morphological rules apply, or from the morphological differences between these elements, rather than from the prosodic domain that is relevant to the various processes at hand.

5.1. Suffixes and enclitics

Glide insertion to break a hiatus exemplifies the distinct behavior of suffixes and enclitics (see also chapter 3, sections 7 and chapter 4, section 3.2). As shown in (326a), a glide is inserted between a stressed /e/ and a following vowel, which in these cases is part of an inflectional affix – a person-number suffix in the first case and the present subjunctive marker in the second (recall that /e/ centralizes to [ɐ] when followed by a heterosyllabic palatal segment). By contrast, the process does not apply if the following vowel belongs to an enclitic, as illustrated in (326b).

- (326) a. *creio* [krɛju] ‘I believe’ (cf. *crer* [krér] ‘to believe’)
 receia [ɾisɛjɐ] ‘he fears’ (cf. *recear* [ɾisjár] ‘to fear’)
- b. *dê-o* [dɛu] ‘give it’
 lê-a [léɐ] ‘read it’

Assuming that this process applies within the prosodic word domain, and given that enclitics are argued to incorporate into the preceding prosodic word, the only factor that may explain this difference is the fact that the process is not operative when the host and the clitic are combined.

Heterosyllabic /e/-centralization shows the same type of distinction between suffixes and enclitics. As we have seen in chapter 3, this process applies when /e/ is followed by a heterosyllabic palatal segment.²⁶³ (327a) shows that it applies if the palatal segment belongs to a derivational suffix. However, it does not operate if the palatal segment belongs to an enclitic, as in (327b) (see also chapter 4, section 3.2).

- | | | | | |
|----------|-------------------|-----|------------------------------|-------------------------|
| (327) a. | <i>plebeizar</i> | [ɐ] | a'. (plebeizar) _ω | 'to turn into plebeian' |
| | <i>europaizar</i> | [ɐ] | (europaizar) _ω | 'to turn into European' |
| b. | <i>dê-lho</i> | [é] | b'. (dê-lho) _ω | 'give it to him' |
| | <i>lê-lhas</i> | [é] | (lê-lhas) _ω | 'read them to him' |

Again, assuming that this process applies within the prosodic word domain and that enclitics incorporate into the preceding prosodic word, the explanation for the distinct behavior should lie in the fact that the process is not operative when the host and the clitic are combined, that is, postlexically. Notice that /e/-centralization does not apply between a prefix and its base, as shown in (328), although the two elements are lexically combined. Here, unlike in the preceding case, the observed behavior follows from the different prosodization of prefixes and suffixes: since prefixes, unlike suffixes, are not incorporated into the prosodic word that dominates their base, they are predicted not to undergo phonological rules that apply within the prosodic word domain. This analysis thus accounts for the fact that heterosyllabic /e/-centralization does not apply between a prefix and its base.

- | | | | |
|-------|-----------------|--------------|--------------------|
| (328) | <i>reisolar</i> | [i]/[j]/*[ɐ] | 'to isolate again' |
| | <i>reimitar</i> | [i]/[j]/*[ɐ] | 'to imitate again' |

5.2. Prefixes and proclitics

Like in the case of suffixes and enclitics, we predict that, given the similar prosodization of prefixes and proclitics, any phonological difference between the two types of elements should derive either from the morphologi-

cal distinction between the two elements, or from the locus of application of particular phonological processes. Final nasal diphthongization exemplifies the distinct behavior of prefixes and proclitics due to morphological differences between these elements. As for the differences between prefixes and proclitics resulting from the locus of operation of pure phonological processes, we will see further below that they do not seem to exist in EP.

As we saw in chapter 3, section 5 there are no nasal diphthongs in non-final stressless positions in the variety of EP under observation here. Thus, we can assume that the prefix *en-* ([ẽ]) and the function word *em* ([ẽj]) are underlyingly similar, if we accept, in addition, that *em* is a stressless word and that *en-* is adjunct to the following prosodic word (like *re-* and proclitic words): both elements should have the form /e[+nasal]/. Under the analysis that final nasal diphthongization refers to the final position of morphological words, rather than to the final position of prosodic words, we are able to explain why the prefix is never realized with a nasal glide, whereas the clitic may be so realized (see 329).

- (329) a. *enviuvar* [ẽ] 'to widow'
 b. *em vitrines* [ẽj] (<[ẽj]) 'in display windows'

Notice that the analysis of this process as referring to the morphological word rather than the prosodic word is corroborated not only by the fact that it applies to the clitic word *em*, but also by the occurrence of diphthongization also at the right edge of a verb form whether or not it is followed by an incorporated enclitic (see chapter 3, section 3.4 and chapter 4, section 3.2).

Another phenomenon that appears to distinguish prefixes from clitics is the reduction that affects clitic function words. This phenomenon, which will be investigated in detail in chapter 7, explains for instance the possibility of function words headed by schwa to be realized without a vowel when followed by a word starting with a vowel. As shown by the contrast in (330), segmentally similar prefixes may not reduce under the same circumstances.

- (330) a. *reanalisar* [j]/*0 'reanalyze'
 reorganizar [j]/*0 'reorganize'
 b. *de analisar* [j]/0 'of analyzing'
 se organizarmos [j]/0 'if (we) organize'

In order to account for this difference, we have previously put forward the hypothesis that this type of reduction is a postlexical phenomenon (since, as it will be shown in chapter 7, it is gradient and sensitive to factors such as speech rate and position within the intonational phrase domain), which operates when proclitics are not yet prosodically attached to their hosts (cf. Vigário 1999a). This would explain why prefixes, which are lexically attached, are not allowed to undergo reduction, unlike clitics.

However, such type of reduction is also shown in chapter 7 to affect very frequent words, including not only clitics but also full prosodic words. Consequently, it is possible that the difference noticed between prefixes and proclitics simply results from a frequency effect. Under this view, it is expected that derived words with the prefix *re-* are not among the most frequent words of EP, unlike clitics that end in schwa. This expectation is in fact corroborated by the data in the EP frequency list of Bacelar, Marques, and Segura da Cruz (1987), since there are no prefixed words with *re-* in the list (see chapter 7 for the ranking of schwa words in this list, as well as for further details on the corpus that it is based on).²⁶⁴

The latter approach to these facts is indeed preferable to the former, since it allows us to dispense with a division within the postlexical component.²⁶⁵

As we have said above, cases where the distinct behavior of prefixes and proclitics results from the locus of operation of pure phonological phenomena appear not to exist in EP. Since pure phonological phenomena apply with reference to prosodic domains, in order to have such a distinction we should have in this language a lexical phenomenon that applies with reference to either the syllable or the foot domain, or that refers to the left-edge of the prosodic word domain. As we have seen in section 4.2, EP does not seem to have phenomena that operate with reference to the syllable or the foot domain at the lexical level. By contrast, there is one process that affects prefixes which refers to the left-edge of the prosodic word, and operates at the lexical level, which is /r/-strengthening (see chapter 3, section 9 and section 3.2 above). Unfortunately, however, there are no clitics that start with /r/. Thus, although we predict that in such cases the clitic should not undergo the rule, we cannot support this claim. As for the phonological processes that apply within prosodic words at the lexical level, they never affect prefixes, since they are adjoined to the prosodic word (see also section 3.2).²⁶⁶ Consequently, we found no distinction between the behavior of prefixes and proclitics that may result from the locus of operation of pure phonological phenomena in EP.

We may thus conclude that the phenomena discussed in this section support the separation of the lexical and postlexical components assumed throughout the present book and is consistent with the analysis defended here.

6. Conclusion: towards the definition of the prosodic word domain in EP

The evidence considered in this chapter leads to the conclusion that different morphological units may be grouped together to form a prosodic word, yielding either incorporated structures (as in the case of suffixed words and host plus enclitic combinations), or adjunction structures (as in the case of prefixed words and proclitic plus host combinations). Our goal in this final section is to determine the general principles that may derive the possible structures observed, and only those.

In order to define the prosodic word domain in EP, we first need to identify which morphosyntactic structure is attributed to derived words, as well as to host plus clitic combinations. As for the morphological structure of derived words, we will adopt here the analysis proposed in Villalva (1994). With respect to the syntactic structure of host plus clitic combinations, we will start by assuming the general approach of Duarte and Matos (2000) for preverbal and postverbal pronominal cliticization.

Essentially, Villalva (1994: 3.3.1) provides the following definitions of morphological constituents: a *stem* (“radical”) is a lexical unit that belongs to a major syntactic category (i.e. there are adjectival, nominal and verbal stems) and that excludes the theme constituent (“constituente temático” – a constituent formed by the theme vowel in the case of verbs and by theme index in the case of nouns and adjectives) (see 331a); the *theme* (“tema”) is a constituent formed by a stem and a theme vowel, in the case of verbs, or, in the case of adjectives and nouns, a theme index (“índice temático”) (see 331b); the *word* is defined as a constituent fully specified that is formed by the adjunction to the theme of the morphological inflection (which, like the theme index, may have no phonological content) (see 331c).

(331) a.	<i>cert</i>	adjectival stem	cf. <i>certo</i>	‘right’
	<i>livr</i>	nominal stem	cf. <i>livro</i>	‘book’
	<i>intruj</i>	verbal stem	cf. <i>intrujar</i>	‘to cheat’

- b. *intruj a* verbal theme (1st conjugation) cf. *intrujar* ‘to cheat’
beb e verbal theme (2nd conjugation) cf. *beber* ‘to drink’
poet a nominal theme (–a) cf. *poeta* ‘poet’
- c. *bebe ssem* word (verbal theme+inflection) cf. *bebessem*
‘they would drink’
poeta s word (nominal theme+inflection) cf. *poetas* ‘poets’

Given these definitions, Villalva (1994: 3.3.2) proposes that EP affixes may attach to different morphological constituents, as listed and exemplified in (332).

- (332) a. derivational suffixes may attach either to stems or to themes
e.g. *certeza* (stem=cert) ‘certainty’
organização (theme=organiza) ‘organization’
- b. inflectional suffixes attach to themes
e.g. *certos* (theme=certo) ‘right(Pl)’
- c. *z*-evaluative suffixes and *–mente* attach to words
e.g. *papelzinho* (word=papel) ‘little sheet’
antigamente (word=antiga) ‘formerly’
- d. prefixes attach to words
e.g. *desfazer* (word=fazer) ‘undo’
subclasse (word=classe) ‘subclass’

If we adopt Villalva’s analysis, we may partially account for the prosodization of both derived and inflected words with the generalization that the prosodic word domain corresponds to the morphological word node that dominates a stem and any adjacent affixes. Indeed, given the definition of *dominance* of Chomsky (1986), already presented in section 3.2, this formulation excludes the morphemes that select a word as their morphological base, which, according to Villalva, are adjoined to the morphological word node. The constituents that select for a *word* are prefixes, as well as *z*-evaluative suffixes and *–mente*. As for the latter elements, which constitute stress domains independent of their morphological base, they will be dealt with in chapter 6. As for stressless prefixes, they require a further remark since we must add the necessary information for these elements to be adjoined to the morphological base they attach to. If we assume that in EP any element that is not dominated by a prosodic word node must be adjoined to the following prosodic word, we may derive both the prosodiza-

tion of stressless prefixes and, as we will see below, the prosodization of proclitic words. The relevant generalization is given in (333).²⁶⁷

(333) Generalization of prosodic parsing at the prosodic word level

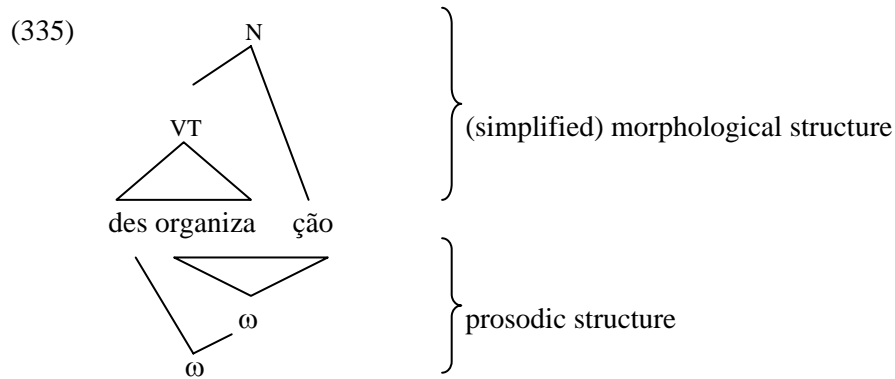
Any morpheme not dominated by the prosodic word node is adjoined to the following prosodic word (if any).

Although prefixes may attach to words, there are also prefixed words where the prefix appears to attach to a theme. This may be illustrated with the example in (334). As stated in Villalva (1994: 49), the prefix *des-* selects either a verbal or an adjectival base. This means that in the nominal form in (334) the prefix must have attached to the verbal theme (VT) before the nominal suffix *-ção* was added to the morphological constituent formed by the prefix and the verbal theme, like in the representation in (334a) and unlike in the representation in (334b) (cf. Villalva 1994: 50).

- (334) *desorganização* (*des* - prefix; *organiza* - verbal theme; *ção* - suffix)
 ‘disorganization’
 a. [[*desorganiza*]_{VT} [*ção*]]
 b. *[*des*] [*organização*]_N

Consequently, here the prefix is not attached to a *word* but rather to a *theme*. Nevertheless, in this case too the prefix behaves phonologically like an adjunct to the prosodic word that includes the prefix morphological base. This constitutes therefore an instance of a mismatch between the morphological structure and the prosodic structure, since, as illustrated in the simplified structure in (335), there is no morphological constituent that excludes the prefix and includes the suffix.²⁶⁸

Given these observations, it is not possible to maintain the proposal that the prosodic word domain corresponds to the morphological word node that dominates a stem and any adjacent affixes, as stated above. The generalization formulated in (336), by contrast, may account for both the cases where prefixes are attached to words and the cases where they attach to themes.²⁶⁹



(336) Lexical prosodic word domain

The prosodic word domain includes a stem plus suffixes
(and non-transparent prefixes)

This definition, together with the generalization that in EP elements not dominated by a prosodic word node are adjoined to the following prosodic word, stated in (333) above, accounts for the prosodization of words at the lexical level.

Let us now consider the information required to derive the prosodic structure assigned to clitics. We will start our discussion with a brief description of the syntactic structure of the sequences comprising pronominal clitics.

The syntax of pronominal clitics is one of the most recurrent topics of investigation in the EP literature (cf. Galves 1992; Madeira 1992; Martins 1994; Rouveret 1999; Duarte and Matos 2000, among many others). As it is clear from the survey of different proposals presented in Duarte and Matos (2000), researchers do not agree on the syntactic position where the clitic ends up with respect to the verbal host. For example, in some of this work, the pronominal clitic and the verb may remain in two distinct functional heads (this is the case of some proclitics in Rouveret 1999, and of enclitics in Martins 1994), whereas in other work, both preverbal and postverbal pronominal clitics are attached to the functional head that includes the verb (as in Duarte and Matos 2000).

We have seen that pronominal enclitics show the maximal degree of cohesion with their hosts. In addition, proclitics and enclitics are differently prosodized with respect to their host. However, this difference cannot be

attributed to the clitics themselves, since the same clitics are differently structured. The difference noted, therefore, may only follow from the position of the clitic relative to the host. This conclusion leads us to assume that at least postverbal pronominal clitics are attached to the same syntactic head that includes the verb, along the lines of Duarte and Matos (2000). The relevance of this assumption for the prosodization of postverbal pronominal clitics will become clear further below.

Pronominal enclitics raise also an interesting question concerning their prosodization, since postlexically inserted clitics are prosodically incorporated into the host prosodic word, and thus the right-edge of the lexical prosodic word (corresponding to the verb) is not preserved at the postlexical level. The question to be answered is how to account for this mismatch between lexical and postlexical prosodic word boundaries. There are at least two ways of obtaining this result. One way consists in admitting the existence of a postlexical restructuring process that erases the right-edge of the lexical prosodic word and inserts a new prosodic word boundary to the right of the enclitic. We exclude this hypothesis both on the basis of simplicity and because a more general and productive analysis is available.

Our proposal is to take advantage of the essential difference between the lexical level and the postlexical level with regard to the definition of the prosodic word. In particular, we will explore the possibility that the postlexical definition of the prosodic word domain includes two components: one that defines the way the lexical prosodic words are related to the postlexically built prosodic structure; and another that defines the syntactic information referred to in the mapping between the syntactic structure and the prosodic structure at the level of the prosodic word.

The first component is required because not all information that comes from the lexicon seems to be preserved postlexically, as we saw above. Given that enclitics are incorporated into the preceding prosodic word, instead of being adjoined to it, we must conclude that the lexical prosodic word boundary to the right of the verb is not projected postlexically. In contrast, the fact that proclitics, as well as prefixes, occur in adjunction structures that are seen postlexically suggests that the left-edges of lexical prosodic words are postlexically projected, as stated in (337) below.²⁷⁰

As for the second component, it is needed, for instance, in order to account for the prosodization of enclitics, since we have to explain the presence of the prosodic word boundary to their right, given that enclitics are not independent prosodic words. If we accept that postverbal pronominal clitics are part of the syntactic head that also includes the verb, we may assume that the mapping relation between syntax and phonology is such that

the lowest syntactic constituent – the X^0 node – always corresponds to the lowest postlexically built prosodic constituent (the prosodic word). This is also stated in (337), where Lex^0 expresses the general property that phonology is not sensitive to empty categories or functional heads that do not include lexical elements (see Truckenbrodt 1999 for a general discussion).²⁷¹

(337) Postlexical prosodic word domain

- a. (Only) The left edges of a lexical prosodic word are postlexically projected.
- b. A Lex^0 is mapped onto a prosodic word.

With respect to the remaining clitics, we may assume that they are subject to the general condition stated in (333) above: morphemes that are not dominated by a prosodic word node are adjoined to the following prosodic word.²⁷² This generalization accounts for the fact that in EP, except for postverbal pronominal clitics (which get enclitic because of the syntactic position they occupy, under the same head as their host), all other clitics are proclitic rather than enclitic. In addition, this condition is extremely general and simple, allowing the integration of clitics into prosodic structure with very little information regarding the syntax associated to these function words.

Finally, we still have to account for the marked situation of words that are lexically unstressed but may exceptionally form independent prosodic words. First of all, this seems to be a marked situation because it only happens when stressless words occur in intonational phrase final position. Interestingly, the fact that these elements in such a position may not become enclitic lends support to the importance of generalization established above that unattached morphemes in EP must be linked to their right. When a stressless function word appears in intonational phrase final position there is no prosodic word to their right. Given the generalization that clitics attach only to the right in EP (except if they are to the right of their host within the limits of its Lex^0), it follows that, in such a position, these function words may not surface as clitics. Consequently, the last resort alternative seems to be their promotion to prosodic words. Since in this position the relevant items become the prominent elements of their intonational phrase, they receive stress and thus the fundamental property that prosodic words must bear stress is satisfied.

We leave the final proposal on the definition of the prosodic word domain in EP to chapter 6. We will see that the general features of the

proposal just outlined may also account for the prosodization of a number of constructions that, although formed by more than a single prosodic word, still behave like (compound) prosodic words. The observation of these constructions is consequently required so that an integrated final proposal regarding the prosodic word domain in EP can be made.

²¹³ For a similar remark on the cues for the prosodic word in English, see Raffelsiefen (1999a). Notice that, as mentioned in section 1.5.1, (re)syllabification is also frequently cited as a useful test for this prosodic domain in Germanic (but not in Romance) languages (e.g. Booij 1995, 1996a; Wiese 1996; Hall 1999a; Raffelsiefen 1999a).

²¹⁴ Notice, in addition, that word stress is assigned with reference to a domain morphologically defined, rather than with reference to the prosodic word domain (see section 3.1).

²¹⁵ This sort of processes may in fact exist in languages that impose minimality restrictions on prosodic words. An example of a rule argued to be motivated by minimality requirements is epenthesis in imperative forms of monosyllabic stems in Shona and KiNande (two bantu languages)—cf. Downing (1999) and references therein. Leben and Ahoua (1997: 126) also suggest that the inclusion of monosyllabic words into adjacent prosodic words in Baule is motivated by minimal word requirements.

²¹⁶ Recall that under the analysis presented in section 3.4 final nasal diphthongization consists of the semivocalization of an underspecified nasal segment in coda position, rather than on glide insertion.

²¹⁷ For this computation we have considered nouns, verbs, adjectives, adverbs, and stressed pronouns, and we have excluded function words—e.g. determiners, prepositions, pronominal clitics, conjunctions, complementizers and other particles—since some function words do not form independent prosodic words, as we will see in section 5.3.1.2.

²¹⁸ Note that the two corpora and methodology are not entirely comparable: we have excluded from the observation function words, which are often monosyllabic, and the words considered are inflected; whereas Thornton (1996) evaluates the proportion of monosyllabic nouns with respect to the whole set of words contained in the Italian Basic Vocabulary, regardless of their category.

²¹⁹ There are two possible realizations of this word, depending on the application of the semivocalization of the vowel preceding the schwa (see section 3.10.2 for the description of this optional process). Naturally, only in the realization where V1 does not semivocalize is schwa in syllable initial position.

²²⁰ See Peperkamp (1997a) on the prosodic word readjustment caused by resyllabification (see also section 1.6 of this book for a review).

²²¹ Although it is generally accepted that there are two sources for schwa in EP (one resulting from the application of vowel reduction to non-back vowels and the other resulting from an insertion process), they are usually not explicitly differentiated. Nevertheless, the presence of schwa has been related to factors like speech style and/or rate (e.g. Mateus et al. 1990: 303).

²²² In contrast with other Germanic languages, English allows for words to begin with a schwa (cf. Raffelsiefen 1999a: 148, among others).

²²³ See Nespor (1985) for a similar observation in Italian. It is also well-known that [ð] in English is only possible word initially with function words (e.g. Morgan, Shi, and Allopena 1996, among many others).

²²⁴ We omit here pitch accent and focus assignment, since it may be thought that these phonological events are associated with prosodic word stressed syllables and thus may constitute primary evidence to word stress, rather than to the prosodic word, directly. See, however, chapter 6, in particular, section 6.2.1, for the suggestion that pitch accent distribution as well as focal stress are sensitive to prominence relations defined within compound prosodic word domains. As for the I-initial H-tone, it usually aligns with the edge of the prosodic word domain (see section 3.17) and thus we will treat it as a boundary tone, which may therefore cue the prosodic word domain independent of word stress.

²²⁵ As we will see in section 6.2.1, in addition to the suffixes identified at least ever since Gonçalves Viana (1883) to bear word stress independently of their base (z-evaluative suffixes and the adverbial suffix *–mente*), we could find another affix-like unit with the same properties, which is *–avos*.

²²⁶ We assume that [r] is part of the inflectional affixes that mark the future and conditional (see the discussion in section 4.3.3). In the examples we use capital letters in underlying to signal the syllable that bears emphatic stress.

²²⁷ In EP, a schwa between consonants is frequently not realized (e.g. Morais Barbosa 1965; Mateus and Delgado Martins 1982), unless it receives initial or emphatic stress. Since when a schwa in this context is possible its deletion is usually also possible, for the sake of simplicity we will not mark the latter realization in the exemplification.

²²⁸ Notice that the realization of initial vowels remains variable despite the occurrence of resyllabification, which causes the prefix final consonant to become the onset of the syllable starting the morphological base. Resyllabification in EP, like in other Romance languages, has a phrasal domain (the intonational phrase) (cf. Vigário 1999a, and Frota 1998 for relevant acoustic data concerning resyllabification of final fricatives in EP).

²²⁹ The form *reiterar* ‘to restate; to emphasize’ is historically prefixed and the prefix is realized with a [v]. We assume that a reanalysis of this word as a simple word has taken place, since *iterar* is not used as an independent word and consequently the meaning of the word is not compositional.

²³⁰ Notice that, according to Chomsky (1986: 7), in a structure where a constituent α is adjoined to a constituent β , β consists of two ‘segments’, and therefore in such configurations β does not dominate α .

²³¹ Note that in the 118 cases of emphatic stress collected, only 3 show stress in an underlying non-initial vowel of a non-prefixed word. In 2 of these, in addition, the

‘stressed’ vowel corresponds to the first *phonetic* vowel of the relevant words (see note 51 below for further details).

²³² For example, syllable degemination may apply in the sequence (*indique*)_o *que terminou o serviço* ‘indicate that you have finished the work’. According to the formulation in (21), the process may apply regardless of the prosodization of the proclitic *que*, which, in theory, may attach to the following prosodic word, or to higher-level prosodic constituents (see section 1.5.3 for the different possibilities of prosodization of clitics). We owe this observation to Lisa Selkirk. We will see in section 5.3.3, nevertheless, that in EP all proclitics are adjoined to the following prosodic word.

²³³ By contrast, emphatic stress seems to be optionally assigned both to the first and to the second vowel of the word *desodorizante*. Like in chapter 3, in this chapter we use the symbol ^x to indicate a very marginal/almost impossible realization.

²³⁴ But see chapter 4 for the discussion on the affix/clitic status of pronominal clitics in EP, and the conclusion that they are not affixes.

²³⁵ Notice, however, that in the case of simple clitics the presence of postlexical stress assignment blocks the occurrence of the clitic (reduced) form, whereas in the case of special clitics and bound words such stress assignment does not prevent the occurrence of the clitic words. Thus, the latter class of clitics may bear (postlexical) stress. Languages reported to have lexically unstressed clitics that may end up with a postlexically assigned stress include Greek (cf. Nespor and Vogel 1986; Berendsen 1986: 4.2; Nespor 1999b), French (cf. Van der Leeuw 1997; Nespor 1999b), Neapolitan (cf. Bafale 1994; Peperkamp 1997a; Nespor 1999b), and Southern Calabrian (cf. Nespor 1999b), among others.

²³⁶ The expression ‘at most’ is intended to allow for clitics to be smaller than a syllable, as in the case of some clitics in English (cf. Selkirk 1984), Dutch (cf. Berendsen 1986; Booij 1996), or German (cf. Wiese 1996; Hall 1999a).

²³⁷ For example, while auxiliary verbs and modals may be clitic in languages like English (e.g. Selkirk 1984, 1996), the elements that belong to these classes in Igbo bear word stress (cf. Zsiga 1992: 128). Similarly, some adverbials are clitic in Dutch (e.g. Booij 1996), but not in EP. Conversely, the definite article is a clitic in EP, whereas determiners in Korean form autonomous prosodic words (e.g. Jun 1996). On the identification of EP clitic words see section 5.3.1.2.

²³⁸ To be more accurate, [v̥] may appear in stressed position but only when followed by a palatal segment (deriving from an underlying /e/)—see section 3.5—or followed by an underlying nasal segment (deriving, in this case from an underlying /a/)—cf. Mateus (1975), among others.

²³⁹ We assume that the obligatorily contracted forms involving prepositions and articles, as well as personal pronouns, are lexically listed as single units (see, section 4.3.2.2 for the analysis of EP personal pronouns, and, in particular, Napoli and Nevis 1987 for the defense of a similar approach to certain preposition plus article clusters in Italian).

²⁴⁰ Clitic words ending in schwa may also undergo further reduction, that is, their only vowel may be deleted (see chapter 7). In addition, with enclitic schwa-words (i.e. postverbal pronouns) vowel deletion is obligatory. This is so because, as argued in section 5.3.3.1 below, enclitics are incorporated into the host's prosodic word and thus become prosodic word final. The incorporated clitics therefore create the environment for the application of the rule that deletes non-back vowels in prosodic word final position.

²⁴¹ To be more accurate, we will see in section 5.3.3.1 that some function words can receive a pitch accent if they are assigned intonational phrase prominence. This seems only to occur with complementizers and relative pronouns in intonational phrase final position (see also Frota and Vigário 1996).

²⁴² To be rigorous, the stressless form *que* may appear either preceded by *o* or not, while the stressed form *quê* must always be preceded by *o* (see Âmbar 1992 for a syntactic account of Wh-movement in EP and of the form of Wh-words in different contexts).

²⁴³ An alternative analysis would be to consider that all interrogative words have a strong and a weak form, although only in some cases we may distinguish these forms on the basis of their phonological make-up. There is, however, an important argument against this hypothesis: the behavior of *porque* and *onde* is distinct with respect to the process of final non-back vowel deletion. In fact, as shown in (i) below, this process usually applies to *onde* but not to *porque*. Notice that in the latter case the vowel may also be deleted, as in other cases of reduced clitic words, but deletion is not obligatory, as it normally happens with prosodic words (see chapter 7 on the distinction between the two processes and on collected data corroborating these observations).

- (i) a. *Porque* abandonaste a sala? [j]/0 'Why did (you) abandon the room?'
 b. *Onde* abandonaste o gato? ^x[j]/0 'Where did (you) abandon the cat?'

This shows that *onde* patterns like a prosodic word, unlike *porque*.

²⁴⁴ We have omitted the cases where a glide precedes the syllable nucleus, because in such environment glides are obtained through the application of V1 semivocalization, which is a postlexical process (cf. section 3.10.2).

²⁴⁵ We should add that EP syllables in full words may also include two consonants in the Rhyme, as in the word *pers.pi.caz* 'acute', although this is extremely rare (see also section 3.7).

²⁴⁶ In fact, some of these syllable structures may arise in reduced forms of clitics or in contractions of clitic clusters, as in (i).

- (i) CCV (para) pra 'for'
 CCVC(para as) prá's 'for-the-FEM-PL'

²⁴⁷ In section 5.3.3.2 below we justify why we do not include final round vowel deletion and final central vowel deletion in the list of processes that can be used to

assess the left-edge of the prosodic word (see section 5.2.2 above on a similar discussion concerning syllable degemination).

²⁴⁸ The judgments of the forms in (46c) are difficult to establish because we have a sequence of a stressed vowel and a high unstressed vowel, and these stressless vowels may be extremely reduced—this is why we have used the symbol ^x to indicate that this realization would be nearly impossible. Given the difference in sonority between the two vowels, for some speakers the second vowel may be felt to become a glide. According to our judgments, however, it is not possible to syllabify the sentences in (46c) as *[éw.víw.tɛj̥f.tu], *[sɔ̃.ɐ.tíw.díj̥f] but only as [éw.ví.u.tɛj̥f.tu], [sɔ̃.ɐ.tí.u.díj̥f] (where the dots signal syllable boundaries). By contrast, both possibilities are available in the examples in (46b): [...ví.u...]/[...víw...] and [...lí.uj̥f...]/[...líwj̥f...].

²⁴⁹ The segmental composition of vowel initial clitics does not allow to test the application of the rule to the final vowel of a verb followed by a postverbal clitic. In fact, vowel initial pronominal clitics have either a central vowel or a high vowel. In such a configuration, as reported in section 3.13, the rule does not apply, independently of the prosodization of clitics. Despite this, it is possible to establish a contrast between the deletion of the central vowel when followed by a high vowel that belongs to a(n en)clitic, which is totally impossible (cf. ia), and the deletion of this vowel when followed by a high vowel that belongs to another full word, since the latter yield, nevertheless, a better result (cf. ib).

- (i) a. (isso,) diga-o você [ɐ]/*0 ‘(as for that,) say-it yourself’
 b. casa usada [ɐ]/*0 ‘used house’

²⁵⁰ Another fact that suggests that this is not a pure phonological process is that in a few specific sequences involving very frequent function words, and in certain familiar registers, a sequence of vowels may also undergo merger to [ɔ] even though the first one is not [ɐ] or [a]. This is documented in (i), where some of the possible realizations are given.

- (i) *todo o lado* [tódu]/[tódɔ] ‘everywhere’
 como o João [kómu]/[kómɔ] ‘like (the) John’
 mais do que o Pedro [dukju]/[duku]/[dukɔ] ‘more than (the) Peter’

This sort of merger is in general impossible with other words (cf. ii).

- (ii) *eu mando o correio* [mẽduu]/*[mẽdɔ] ‘I send the mail’
 eu como o gelado [kómu]/*[kómɔ] ‘I eat the ice-cream’
 indique o caminho [ĩdíku]/*[ĩdíkɔ] ‘show the way’

The phenomenon is possibly conditioned by word frequency (see chapter 7 for other phenomena conditioned by word frequency).

²⁵¹ In the case of *ao* we will consider the monophthongized form to be the result of a reduction process in chapter 7. It is possible that the same line of explanation may be pursued in other cases where clitics and perhaps other frequent words are involved.

²⁵² As we have seen in section 5.3.1, the function word *pelo* may have reduced forms. In order to assess the application of syllable degemination it is important to select the unreduced form [pelu], since the reduced form [plu] shows a complex onset, and therefore does not provide the segmental context for the rule to apply (see section 3.14 for the details concerning the segmental conditions on this process).

²⁵³ Given the existence of solid arguments for the incorporation of postverbal pronominal clitics into the verb's prosodic word, and since enclitics do not affect word stress location, we must conclude that word stress location is not a valid argument for determining the prosodization of clitics in EP. By assuming that word stress is a lexical rule (see section 3.1), we have actually predicted that the postlexical combination of clitics with their hosts could not affect word stress. In chapter 8, we review similar data of Standard Italian in the light of EP facts and discuss the consequence of our findings for the prosodization of enclitics in that language.

²⁵⁴ In fact, no sentence in EP may end in any of the function words considered in this chapter (to the exception of the form *cada* 'each' when it is used as a pronoun). In general, only complementizers may end an intonational phrase. Only when a parenthetical expression is inserted after a preposition (which is always rather marginal), or after a relative pronoun may such a function word occur in this configuration. In the case of the definite article this seems totally impossible. The reasons for this are possibly syntactic rather than prosodic (since, for example, demonstrative items like *este* 'this' are stressed, but may not be separated from their noun phrase either). These peculiarities fall outside the scope of the present work, and thus we will not develop them further.

²⁵⁵ This assertion is based on informal observations of uttered sentences by EP speakers, as well as on introspection. Although we believe this description does not raise any controversy, to our knowledge this question was never raised in EP literature and thus we found no reference to this distinction between complementizers and postverbal weak pronouns.

²⁵⁶ According to Bafale (1994), in Neapolitan verb-clitic sequences with two stresses, the first is a secondary stress, while the stress on the clitic is the primary one.

²⁵⁷ We will not explore here other possibilities for the phrasing of these sentences. It is however important to observe that alternative (preferred) prosodizations consist of the integration of the complementizers either into the following or into the preceding I-phrase, as already described in Frota and Vigário (1996) and Vigário and Frota (1998). The existence of the three possibilities may be due to the interaction between the mapping algorithms of I-construction, on the one hand, and

the prosodic weight requirements on I-phrases, on the other hand (on these issues, see Frota 2000a and section 1.1).

²⁵⁸ Notice that even under this postlexically assigned stress, the final vowel of *que* and *porque* surfaces as a schwa. In fact, these complementizers have no strong variant, unlike the homophone interrogative pronouns (see section 5.3.1.2).

²⁵⁹ The present discussion excludes the cases where a stressless word follows a stressed function word (like the emphatic expression *é que*, and *se bem que/já que/só que* ‘although’, among others). In fact, in these cases, it is possible that the cluster of function elements, which appear together very frequently, may have lexicalized as a single unit. These cases deserve a specific study, and thus are left for future investigation.

²⁶⁰ This behavior is reminiscent of the behavior of the Šrem dialect of Serbo-Croatian described in Selkirk (1996), where proclitics are assigned the same representation as the one we are proposing for EP (see section 1.5.3).

²⁶¹ Gonçalves Viana (1883: 62) also rules-out the possibility of a *secondary stress* assigned to postverbal personal pronouns in EP.

²⁶² Of the 118 attested tokens with emphatic stress in Appendix I, in 110 emphatic stress occurs in the first syllable of either a full word or of a clitic; in 5 cases it occurs in the second syllable of a prefixed word, which also counts as prosodic word initial (see section 5.2.2); in only three cases is the emphatic stress realized on a non-initial syllable. Notice that, even in the latter cases, it occurs in the first phonetic vowel in two tokens, and on the second phonetic vowel of the word in one token. These cases do not therefore constitute important counter-evidence to our proposals, which account straightforwardly for 97.5% of the collected data.

²⁶³ As noticed in section 3.5, heterosyllabic /e/-centralization does not operate when this vowel is in a stressless position and is followed by a palatal consonant. We have assumed that this is so because vowel reduction (i.e. raising and centralization) has operated. When the vowel is followed by a high palatal vowel, by contrast, vowel reduction is regularly blocked, and therefore the conditions for centralization to apply are met. There are no suffixes that start with a palatal consonant and are preceded by a stressed /e/, and thus the forms in (88) are not maximally comparable. We recall, nevertheless, that in simple words, heterosyllabic /e/-centralization regularly applies in forms that are segmentally nearly identical to those involving clitics in (88b) above (cf. *tenho* [tɛ̃nu] ‘I have’, *telhas* [tɛ̃ʎɐʃ] ‘tiles’—see also section 4.2.2).

²⁶⁴ To be rigorous, there is one word in the list cited above that is historically derived with the prefix *re-* which is *reunião* [ʀjunjɐ̃w] ‘meeting’ (see also section 5.2.2 on the phonology and semantic interpretation of this word form). Interestingly, as we will see in chapter 7, this particular word is sometimes realized with a very reduced glide, which, in extreme cases, may even be absent. This supports the present approach that glide deletion is correlated with the frequency of

words rather than with the lexical/postlexical locus of prosodization of prefixes and proclitics.

²⁶⁵ We thank Geert Booij for drawing our attention to this point.

²⁶⁶ By contrast, prefixes may undergo postlexical span rules that have a domain higher than the prosodic word (e.g. V1 semivocalization), as well as postlexical phenomena that refer to the left-edge of the prosodic word domain (e.g. initial stress assignment) (see section 1.1, note 12 the notions of span, limit and juncture rules). However, as expected, such phenomena also affect proclitics.

²⁶⁷ The expression “if any” in (94) is intended to cover the cases where there is no following prosodic word. This happens with clitics at the lexical component, since they are only combined with their hosts postlexically, as well as with proclitics at the postlexical component when they appear in intonational phrase final position (see section 5.3.3.1).

²⁶⁸ From the point of view of the morphological structure, according to Villalva’s (1994) approach, *-ção* is added to the verbal theme, yielding a nominal *theme*, headed by a theme index; to the theme the morphological inflection (phonetically empty in this case) is then attached, which finally creates the *word* of the nominal category.

²⁶⁹ Other languages where a stem plus any suffix are grouped together under the same prosodic word include Italian (Nespor and Vogel 1986; Peperkamp 1997a), French (Hannahs 1995b), and Spanish (Peperkamp 1997a).

²⁷⁰ See, in particular, Selkirk (1986, 1996) for an edge-based approach to prosodic domains construction.

²⁷¹ The generalization that *Lex^o* is mapped onto a prosodic word will also play a major role in the prosodization of compound-like constructions (see section 6.4). A similar proposal has been put forward in Nespor and Ralli (1996) and Nespor (1999b) to account for the prosodization of compounds in several languages, although the consequences extracted from it are not precisely the same as ours (see section 6.4 and, for a review of this proposal, section 1.5.4).

²⁷² Notice that, under the present view, pronominal proclitics end up adjoined to the verb’s prosodic word either as a consequence of the general condition of rightwards adjunction, or as a consequence of the definition of the postlexical prosodic word: (i) if they are not part of the same *Lex^o* as the verbal host, they are prosodized like other clitic function words; (ii) if the proclitic is part of the syntactic head that includes the verb instead, the prosodic word boundary to the left of the verb is obtained because it comes from the lexical level (and left boundaries are postlexically projected), and the prosodic word boundary to the left of the clitic follows from the mapping of *Lex^o* nodes onto prosodic words.

Chapter 6

Compound Prosodic Words

1. Introduction

In this chapter we investigate the prosodization of words (e.g. specific derived words) and combinations of words (e.g. some compounds) that exhibit more than a single word stress, and whose phonological behavior will be shown to be distinct from the behavior of words regularly combined within syntactic phrases. Our main goal is to identify the prosodic constituent formed by the grouping of the different stressed units that compose these constructions.

Assuming that there is a one-to-one correspondence between a stress domain and a prosodic word, and since the prosodic constituent above the prosodic word level is the phonological phrase, it could be expected that prosodic words are always grouped together under the phonological phrase domain (as proposed in Nespor and Ralli 1996 and Nespor 1999b; see chapter 1, section 6.4 for a review). Nevertheless, given the relaxation of the Strict Layer Hypothesis argued for by various authors (see chapter 1, section 5), it can also be hypothesized that certain types of stressed units form together a *compound prosodic word* domain (as implicitly assumed in Leben and Ahoua 1997; see also chapter 1, section 6.4 of this book).

In order to decide which of the two possibilities mentioned above holds in EP, we will first review the phenomena assumed to characterize each stress domain, and the phenomena expected to occur across prosodic words (section 2). In the following section, we examine the phonological behavior associated with a number of constructions consisting of more than one stress domain that will be shown to display specific phonological properties. These constructions include (i) derived words with suffixes that constitute stress domains independent of their morphological base (section 3.1), (ii) derived words with stressed prefixes (section 3.2), (iii) stem plus stem compounds (section 3.3), (iv) word compounds (section 3.4), (v) abbreviations (section 3.5), and (vi) mesoclititic structures (section 3.6). Other specific combinations of words that display a particular behavior are also investigated in section 3.7.

Two phenomena that are not purely phonological but that appear to refer to prosodic constituency – namely, deletion under identity and clipping –

are investigated in section 4. It will be suggested that these phenomena support the prosodic analysis proposed throughout this book on the basis of pure phonological criteria.

We conclude this chapter with a discussion on the relation between word stress and the prosodic word domain. It will be shown that not only morphosyntactic information, but also aspects related to word stress play an important role in the prosodization of the structures studied in this chapter. This will lead us to our final proposal on the definition of the prosodic word domain in EP, which integrates the results of the study developed both in the present chapter and in the previous one (section 5).

2. Phonological characterization of prosodic words and their concatenation

In this section, we briefly review the phonological phenomena that can be used as diagnostics for the prosodic word status of a given unit in EP. Additionally, we summarize the processes that regularly apply across but not inside of words. These phenomena will be crucial in the ensuing sections to determine the prosodic structure associated to the constructions under investigation.

As seen in chapter 5, section 2, a number of tests constitute reliable cues for the prosodic word domain in EP. These are listed in (338).

(338) Tests for the prosodic word in EP

- word stress assignment
- phonological phenomena that refer to the prosodic word domain
- phonotactic generalizations

Word stress in EP is usually considered to apply within a morphological domain (e.g. Mateus 1983, d'Andrade 1988, 1997, Pereira 1999, and chapter 3, section 2). Nevertheless, the presence of one primary stress associated to a given prosodic unit implies the prosodic word status of that unit (see chapter 1, section 6.1). Consequently, the presence of word stress identifies the presence of a prosodic word.

The application of some phonological processes depends on the presence or absence of word stress. For example, a vowel bearing word stress may not undergo vowel reduction, semivocalization, vowel deletion, and vowel lowering in syllables closed by /r/ (see chapter 3, sections 3, 11,

12–14, 8, respectively). Conversely, some processes may only apply in stressed environments, as for example glide insertion to break a hiatus, and pitch accents and focal stress are only assigned to stressed positions (see chapter 3, sections 7, 18 and 19, respectively). Thus, all these phenomena may cue the presence/absence of word stress.

Some of the phonological events mentioned in the preceding paragraph constitute also a direct cue for the prosodic word domain, since they occur within or at the edges of this prosodic constituent. Table 2 in chapter 5 summarizes the phonological phenomena that refer to the prosodic word domain in EP (see also chapter 3). We repeat the relevant information under (339) below.

(339) Phonological phenomena that refer to the prosodic word domain

Heterosyllabic /e/-centralization
 Glide insertion to break a hiatus
 Lowering of stressless vowels in final syllables closed by /r/
 Initial /r/-strengthening
 Initial vowel realization (neutralization/feature specification)
 V2 semivocalization
 Final non-back vowel deletion
 Final round vowel deletion
 Final central vowel deletion
 Syllable degemination
 Initial stress
 Emphatic stress
 Tonal association

Among the processes listed, some apply only across words: namely, final round vowel deletion, final central vowel deletion, and syllable degemination. Round vowel deletion and central vowel deletion have been reported, furthermore, to be sensitive to stress-clash configurations, involving ϕ -level prominence (cf. Frota 2000, and chapter 3, sections 13 and 14). These phenomena, therefore, may be useful to contrast the behavior of regular word plus word combinations within ϕ with the behavior displayed by the members of the structures under investigation.

We will see, additionally, that non-back vowel deletion, focus assignment and pitch accent distribution provide important evidence for the specific prosodization of the constructions studied here, which contrast with word plus word combinations in general.

As for phonotactic restrictions, they will play no crucial role in the investigation developed in this chapter. Given that we could only find phonotactic restrictions imposed upon lexical prosodic words, these restrictions are of no use when we consider units that include more than one prosodic word.

Two other processes that may refer to the prosodic word domain, on the one hand, and that involve more than a single prosodic word, on the other, are those mentioned in (340).

(340) Additional tests for the prosodic word domain in EP

- deletion under identity
- clipping

Deletion under identity in other languages has been described to apply with reference to the prosodic word domain in compound and derived words with more than a single word stress (cf. Booij 1985, 1995, Wiese 1996, and chapter 1, section 6.1 for a review). It may consequently cue the prosodic word constituent also in EP. Given that this is not a purely phonological phenomenon, it will be investigated in a separate section, together with clipping (section 6.3). In the latter case, the minimal prosodic word has been proposed to function as a template to which an expression is reduced in languages like Italian (cf. Thornton 1996, reviewed in chapter 1, section 6.1). Although EP prosodic words do not seem to have a minimal size, the prosodic word may, nevertheless, play a role in clipping. Indeed, it will be seen that the prosodic configuration associated to the constructions that undergo both clipping and deletion under identity corroborates the prosodic analysis already supported by the phonological data presented in chapter 6, section 3 and in chapter 5.

3. Prosodization of words with more than one stress domain

Resorting to the phenomena surveyed in the preceding section, we will investigate now the prosodization of various constructions with more than a single word stress, which, in general, are classified morphologically and/or syntactically as *words*.²⁷³

3.1. Suffixed words with more than one word stress

Words that are traditionally assumed to be formed with suffix-like elements that constitute a stress domain independent of their base are those with *z*-evaluative suffixes and the adverbs formed with *–mente* (Gonçalves Viana 1883, Morais Barbosa 1965, Mateus 1983, Villalva 1994, Mateus and d’Andrade 2000, among others – see the review in chapter 2).²⁷⁴

It is uncontroversial that both these suffix-like elements and their base are assigned word stress.²⁷⁵ The first cue for this is the perception of two stresses. A second piece of evidence for the presence of the stress on the morphological base is the regular non-application of vowel reduction to the base stressed vowel, as illustrated in (341).

(341) a.	<i>jacarÉ</i>	[ɛ]	<i>jacarEzInho</i>	[ɛ]/*[ɛ]/*[i]
	‘aligator’		aligator-DIM	
	<i>avÔ</i>	[o]	<i>avOzIto</i>	[o]/*[ɔ]/*[u]
	‘grandfather’		grandfather-DIM	
	<i>trenÓ</i>	[ɔ]	<i>trenOzÃo</i>	[ɔ]/*[o]/*[u]
	‘sled’		sled-AUG	
b.	<i>alEgre</i>	[ɛ]	<i>alEgremENte</i>	[ɛ]/*[ɛ]/*[i]
	‘joyful’		‘joyfully’	
	<i>Ávida</i>	[a]	<i>AvidamENte</i>	[a]/*[ə]
	‘eager-fem’		‘eagerly’	
	<i>pObre</i>	[ɔ]	<i>pObremENte</i>	[ɔ]/*[o]/*[u]
	‘poor’		‘poorly’	

Phonological criteria indicate that there is another morphological unit that patterns like *z*-evaluative suffixes and *–mente*, which is *avos* ‘-th’. Before proceeding, however, a parenthesis on the morphological status of *avos* is in order. Some dictionaries do not include this unit either as a word or as a suffix that forms a word together with a morphological base (e.g. *Universal*, *Dicionário Inverso da Língua Portuguesa*). In other dictionaries (*Lello*, *Lexicoteca* and *Novo Michaelis*) *avo* is classified as a noun, whose meaning is dependent on the numeral it appears with (“the N part of a whole”, where N stands for some numeral) or refers the dominator of a fraction (e.g. *um treze avos* ‘one thirteenth’). However, there are some reasons for us not to treat *avos* like other nouns. First, unlike regular nouns, *avos* may only appear with numerals. The fact that it occurs only with

words of a specific category is a property *avos* shares with affixes and not with other words in general. Additionally, *avos* may not occur with all numerals, given that when N is a numeral below 11 a single non-derived word is selected (e.g. *meio* ‘half’, *terço* ‘third’, *quarto* ‘quarter’, *quinto* ‘fifth’, and so on). Also pointing to its suffix nature is the fact that the numeral with which *avos* occurs may not be the focus of a Wh-word (e.g. **Em que avos de final o João joga?* ‘In which part of the final John plays?’) unlike the numerals that occur for instance with *parte* ‘part’, which is semantically close to *avos* (e.g. *Em quantas partes o concurso está dividido?* ‘In how many parts is the contest divided?’). We thus assume that *avos* has lost the status of a regular noun, similar to *-mente*.²⁷⁶

Returning to the phonological behavior of *avos*, besides the perception of stress on the morphological base – in addition to the stress on the suffix itself – the regular absence of vowel reduction in the stressed vowel of the morphological base shows that this constituent is also assigned word-stress (see 342).

(342) <i>trEze</i>	[e]	<i>um trEze Avos</i>	[e]/*[i]
‘thirteen’		‘one thirteenth’	
<i>dezassEte</i>	[ɛ]	<i>dois dezassEte Avos</i>	[ɛ]/*[e]/*[i]
‘seventeen’		‘two seventeenth’s’	
<i>dOze</i>	[o]	<i>três dOze Avos</i>	[o]/*[ɔ]/*[u]
‘twelve’		‘three twelfth’s’	

In addition to the phonological behavior noted above, *avos* shares with the aforementioned suffixes (and with no other suffixes identified so far) the type of morphological base it selects: a morphological word, in the sense of Villalva (1994) (see also chapter 5, section 6). In fact, all these suffix-like units are added to a base that includes the theme index and inflection.²⁷⁷ This is shown in (343).

(343) a. <i>azul/azuis</i>	<i>azuizinhos</i>
blue/blue-PL	blue-PL-DIM-PL
<i>cão/cães</i>	<i>cãezitos</i>
dog/dog-PL	dog-PL-DIM-PL
b. <i>belo/bela</i>	<i>belamente</i>
beautiful-MASC/beautiful-FEM	‘beautifully’
<i>triste</i>	<i>tristemente</i>
‘sad’	‘sadly’

c. <i>vinte e cinco</i>	<i>um vinte e cinco avos</i>
‘twenty five’	‘one twenty fifth’
<i>trinta</i>	<i>um trinta avos</i>
‘thirty’	‘one thirtieth’

The prosodic word status of the morphological base these suffixes attach to is also supported by the application of a process that occurs at the right-edge of the prosodic word domain: the lowering of stressless vowels in syllables closed by /r/ in prosodic word final position, as illustrated in (344) (see also chapter 3, section 8).²⁷⁸

(344) <i>sEniorzInho</i>	[ɔ]/*[u]
senior-DIM	
<i>lIderzEco</i>	[ɛ]/*[i]
leader-DIM	
<i>IMparmENte</i>	[a]/*[ɐ]
‘uniquely’	

The presence of word stress assigned to the suffix can also be demonstrated by the regular absence of vowel reduction in the suffix stressed vowel (see 345).

(345) a. <i>trenOzEco</i>	[ɛ]/*[i]	sled-DIM
b. <i>dOzE Avos</i>	[a]/*[ɐ]	‘twelfth’

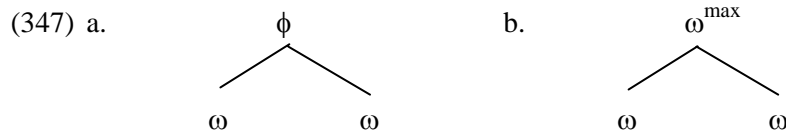
These facts show that both the suffix and its morphological base are assigned word stress. We can therefore conclude that each of these units constitutes an independent prosodic word, as shown in (346).

(346) (trenO) _ω (zEco) _ω
(alEgre) _ω (mENte) _ω
(dOze) _ω (Avos) _ω

To sum up, the suffixes surveyed in this section form prosodic words independent of their base. We should thus point out that EP invalidates Hannahs’s (1995a: 22, note 12) suggestion that, contrary to Germanic languages, in (all) Romance languages suffixes may not form independent prosodic words.

222 *Compound prosodic words*

Let us now investigate the nature of the prosodic constituent formed by the two prosodic words. The two possible representations for these constructions are shown in (347).



Several phenomena point to the prosodization of these constructions as in (347b), rather than as in (347a). The first we will address is focus assignment.

While the possibility of assigning focus to a ϕ -internal prosodic word is described in Frota (2000), and illustrated here in (348), within the complex words under investigation focal stress may only occur in the rightmost prosodic word of the construction, as shown in (349).²⁷⁹

- (348) A: *Foi a tarde âmbar que o pintor retratou?*
 ‘Was it the amber afternoon that the painter painted?’
 B: *Não. O pintou retratou ((a (maNHÃ)_ω)_ω (âmbar)_ω)_ϕ*
 no the painter painted the morning amber
 ‘No. The painter painted the amber morning.’

- (349) A: *Acho que o João trouxe o gato para casa.*
 ‘(I) think that John brought the cat (to) home.’
 B: *Não trouxe não. O João trouxe o cãoZinho para casa.*
 B’: **Não trouxe não. O João trouxe o CÃOzinho para casa.*
 no (he) brought no the John brought the dog-DIM to home
 ‘No he did not. John brought the little dog home.’

Notice that, according to Villalva (1994), *z*-evaluative suffixes are morphological modifiers rather than the nuclei of the structure they appear in (as they may not change the syntactic category of the base or specify morphosyntactic categories). Structure-wise, they are analyzed as adjuncts to the morphological base. Consequently, the association of focal stress to the stressed unit that corresponds to the suffix appears not to be explained by the morphological structure of the word, that is focal stress is not assigned to the morphological nucleus. The hypothesis we put forward is that this association follows from the phonological structure instead. If these complex words are accepted to form a compound prosodic word, we can ac-

Our account of the data just presented is the following: except in intonational phrase initial position (where the initial prosodic word of the compound may be assigned a pitch accent because it is also the initial prosodic word of the intonational phrase) pitch accent assignment to *–mente* adverbs, and arguably other prosodic compounds, looks for the head of the prosodic compound.²⁸² This hypothesis is parallel to that formulated in Vigário (1998) and Frota (2000) to account for the distribution of pitch accents within the phonological phrase (see chapter 3, section 18). Within a prosodic compound the two prosodic words are not treated equally by pitch accent assignment, in the same way as a ϕ -head and a non-head are also distinct from that point of view. This strongly suggests a prosodic structure in which the second (the last) element is the prominent one, both at the level of ϕ and at the level of the compound prosodic word.

Final non-back vowel deletion provides another important argument for compound prosodic words. As we have seen in chapter 3, section 11, prosodic word final non-back non-high vowels are regularly deleted, regardless of what follows. This is illustrated in (352): in (352a) the relevant prosodic word is in sentence final position and its final non-back vowel is deleted; the same is true if the word is followed by a word starting with a consonant (352b), or starting with a stressless vowel (352c); finally, vowel deletion also occurs if the following word starts with a stressed vowel (352d).

- (352) a. *Ela pode* ‘she can’
 b. *Pede tudo à Maria* ‘ask Mary everything’
 c. *Uma suave atmosfera* ‘a soft atmosphere’
 d. *A doce água do rio* ‘the sweet water of the river’

The crucial case for the present discussion is the one in (352d). Unlike in the sequences in (352d) and (353a) below, within the complex words under analysis non-back vowel deletion is regularly blocked if the target vowel is followed by a word starting with a stressed vowel, as in (353b).

- | | | | | | |
|----------|---------------------|---------------------|----|-------------------|--------|
| (353) a. | <i>dOce Água</i> | 0 ^x /[j] | b. | <i>Onze Avos</i> | *0/[j] |
| | ‘sweet water’ | | | ‘eleventh’ | |
| | <i>trIste Época</i> | 0 ^x /[j] | | <i>dOze Avos</i> | *0/[j] |
| | ‘sad period’ | | | ‘twelfth’ | |
| | <i>enOrme Ódio</i> | 0 ^x /[j] | | <i>trEze Avos</i> | *0/[j] |
| | ‘tremendous hate’ | | | ‘thirteenth’ | |

As we will see throughout this chapter, this behavior is by no means specific to this type of word, and therefore a general explanation is called for.

Let us assume that the words in (353b) form a compound prosodic word, and that the phonological process of vowel deletion applies at the right-edge of the prosodic word domain within the domain of a maximal prosodic word (see the definition in 354 below). The context for the blocking of non-back vowel deletion may then be defined as follows: in a configuration $(\dots V1)_{\omega 1} (V2 \dots)_{\omega 2}$, V1 deletion is blocked when V2 bears the stress of the prosodic constituent that dominates both $\omega 1$ and $\omega 2$ – in this case, when V2 bears the compound prosodic word prominence. The same general configuration defines the blocking of other vowel deletion processes (central vowel deletion and round vowel deletion) within the phonological phrase. These processes are blocked when the vowel to be deleted belongs to a ϕ -internal prosodic word and is followed by a prosodic word starting with a stressed vowel which is, in addition, the head of the phonological phrase (see Frota 2000, and chapter 3, sections 13 and 14 of this book). Notice, nevertheless, that these cases are distinct from non-back vowel deletion in that the latter is not blocked when two prosodic words are immediately dominated by the ϕ -phrase, as in (353a).

We propose, therefore, that non-back vowel deletion applies within the maximal prosodic word domain (ω^{\max}), and that it is blocked when the target vowel is followed by the vowel bearing the phonological compound stress.²⁸³ Notice that by assuming the definitions in (354), when the *minimal* prosodic word and the *maximal* prosodic word coincide, there is no *compound* prosodic word. Nevertheless, in such cases, we still have a maximal prosodic word. Thus, even if there is no compound prosodic domain, there is a maximal prosodic word domain where the final non-back vowel of a prosodic word is deleted, regardless of the right adjacent context (as in the examples in 352 above).

- (354) *Compound prosodic word*: a prosodic word that dominates two (or more) constituents of the same type (i.e. the prosodic word);

Maximal prosodic word: a prosodic word that is immediately dominated by the next higher prosodic level (i.e. the phonological phrase);

*Minimal prosodic word:*²⁸⁴ a prosodic word that immediately dominates the next lower prosodic level (i.e. the foot).

Given the preceding discussion, we may now add to the formulation of non-back vowel deletion presented in chapter 3, section 12 the domain within which it operates:

(355) Non-back vowel deletion (reformulated)

$$\begin{pmatrix} V \\ \text{-high} \\ \text{-back} \end{pmatrix} \rightarrow \emptyset / [\dots [\dots _]_{\omega} \dots]_{\omega}^{\max}$$

Round vowel deletion also appears to be sensitive to compound prosodic word prominence. Recall that this process is blocked if the vowel to be deleted (V1) is followed by a stressed vowel (V2) that belongs to the head of a ϕ -phrase (cf. Frota 2000). Vowel deletion is not completely impossible, however, if V2 does not bear ϕ -prominence (see the discussion in chapter 3, section 13). When V2 bears compound stress, by contrast, round vowel deletion becomes impossible, regardless of ϕ -prominence.²⁸⁵ This difference between a sequence of prosodic words within a compound and a sequence of prosodic words within a phonological phrase is illustrated in (356): vowel deletion within the compound is impossible (356a), contrary to vowel deletion within a ϕ -phrase (356b).

- (356) a. (os vinte e quatro avos disputados) $_{\phi}$ *0/[w]
 ‘the twenty fourth (part of a contest) played’
 (os vinte e cinco avos disputados) $_{\phi}$ *0/[w]
 ‘the twenty fifth part (of a contest) played’
 b. vi (quatro aves presas) $_{\phi}$?0/[w] ‘(I) saw four birds caged’
 ouvi (cinco águias selvagens) $_{\phi}$?0/[w] ‘(I) heard five wild eagles’

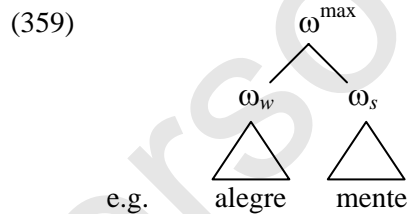
To conclude, we should add that the compound final segments also behave like prosodic word final. For instance, they undergo non-back vowel deletion (357a), and, optionally, round and central vowel deletion (357b–c), as well as syllable degemination (357d).

- (357) a. *somente ofereceram* $^x[j]/0$ '(they) only offered'
 b. *um cãozinho amigo* $[w]/0$ a friendly dog-DIM
 c. *uma avezinha observada* $[v]/0$ a bird-DIM watched
 d. *eles somente temiam uma coisa* $[ti\tilde{t}i]/[ti]$ '(they) only fear one thing'

Similarly, the initial position of these prosodic compounds may be assigned emphatic stress (358a), and the realization of non-central vowels may vary just like in non-compound words (358b).²⁸⁶

- (358) a. *ele é um SOcialIstazInho* he is a socialist-DIM
ele descreveu-a SAtisfatOriamENte 'he described her satisfactorily'
ele disputou os CAtOrze Avos de final 'he played the fourteenth of the final'
 b. *ele está elegantezinho* $[e]/[i]$ he is elegant-DIM
ele vai ocupar $[o]/[ɔ]$ 'he will occupy'

To sum up so far, the data support the compound prosodic word analysis of the constructions with *z*-evaluative suffixes and *avos*, as well as of adverbs formed with *-mente*. In all these cases there are two prosodic words which are grouped together within a constituent that is not the next higher level constituent in the prosodic hierarchy (i.e. the ϕ -phrase), but is rather a constituent of the same type as the constituents it dominates. Focus assignment, pitch accent distribution, and the phonological processes of non-back vowel deletion and round vowel deletion were shown to crucially bear on this issue. The structure in (359) depicts the prosodization proposed for these constructions.²⁸⁷



Before proceeding with our description, it should be noticed that the similarity between prefixes and stems often makes it difficult to set these classes apart (e.g. Cunha and Cintra 1984: 113–115, Villalva 1994: 350; similar difficulties are known to exist in other languages – see, for example, Scalise 1984 and Peperkamp 1997a: chap.4 for Italian, and Bauer 1987: 2.1.6.1 and 5.3.8, for French, German and English). Given the similarity between the two types of units also from a phonological point of view, in our discussion we will consider examples of both stressed prefixes and units that may be classified as prefix-like in EP.²⁸⁸

(360) a. **pré**-acentual [prɛ]/*[prj] 'pre-stressed'
 b. **pró**-comunista [prɔ]/*[pru] 'procommunist'
 c. **pós**-sintáctico [pɔʃ]/*[puʃ] 'post-syntactic'
 d. **bi**-anual [bi]/*[bj] 'biennial'
 e. **ex**-trabalhador [ɐjʃ]/*[ʃ] 'ex-worker'

The contrast between the phonological shape of these prefixes in transparent constructions and their shape in non-transparent constructions, shown in (361), is noteworthy: in examples (b), (d) and (f), vowel reduction

obligatorily applies to the vowel of the prefix; in (a) and (c) V1 semivocalization is possible; and in (e) the initial diphthong is no longer possible.²⁸⁹

- | | | | |
|---------------------------|-------------|--------------------|-------|
| (361) a. <i>preâmbulo</i> | [pri]/[prj] | d. <i>pospôr</i> | [puʃ] |
| ‘preface’ | | ‘to postpone’ | |
| b. <i>promover</i> | [pru] | e. <i>exportar</i> | [ʃ] |
| ‘to promote’ | | ‘to export’ | |
| c. <i>biênio</i> | [bi]/[bj] | f. <i>prever</i> | [pri] |
| ‘two years’ | | ‘to foresee’ | |

The data further show that the realization of the prefix is independent of the segmental context to its right. In fact, stressed prefixes may be followed by bases starting with vowel or with consonant, regardless of ending with a closed or open syllable themselves (e.g. *pós-sintático* ‘post-syntactic’, *pós-operatório* ‘post-surgical’).²⁹⁰

For the sake of completeness, we should add that the presence of word-stress on the morphological base is also clear: it is perceived as such; the stressed syllable does not undergo the processes known to apply to unstressed vowels; and, if the stressed vowel is /e/ and is followed by another vowel, glide insertion applies, as expected in stressed environments (e.g. *pré-recreio* ‘before a break’). In addition to this, the morphological base behaves like an independent prosodic word with respect to other phonological processes. If it starts with a rhotic, it undergoes /r/-strengthening, and initial stressless vowels show the variation in height that is common in prosodic word initial position (e.g. *pré-[e]/[i]mocional* ‘pre-emotional’). Consequently, the internal prosodic structure of these words must take the form shown in (362).

- (362) (pré)_ω (lexical)_ω
 (pós)_ω (operatório)_ω

Let us now consider words with disyllabic prefixes, like those illustrated in (363).

- | | | |
|---------------------------|----------------------|---------------------------|
| (363) <i>poli-copiado</i> | <i>extra-magro</i> | <i>super-interessante</i> |
| ‘copied several times’ | ‘very thin’ | ‘very interesting’ |
| <i>multi-racial</i> | <i>ultra-radical</i> | <i>inter-cultural</i> |
| ‘multi-racial’ | ‘extremely radical’ | ‘intercultural’ |

230 *Compound prosodic words*

mini-aspirador *mega-concerto* *hiper-contente*
 ‘mini-vacuum-cleaner’ ‘huge concert’ ‘tremendously happy’

Similar to monosyllabic prefixes and besides the perception of stress, several facts support the stressed status and the prosodic word nature of both the disyllabic prefix and the base. As for the prefix, the stressed vowel does not undergo vowel reduction, as shown by the forms in (364a). Furthermore, the stressless vowel of prefixes that end in /r/ undergoes the lowering rule that was shown in chapter 3, section 8 to apply at the right-edge of a prosodic word (364b). Finally, the initial diphthong of the prefix *extra-* is not deleted (364c).

- (364) a. *poli-* [póli] b. *super-* [súper] c. *extra-* [éj]trɐ]/*[ʃtrɐ]
 mega- [méɣɐ] *inter-* [ĩ tɛr]

Like monosyllabic prefixes (see the examples in 361 above and also chapter 5, section 3.1), disyllabic prefixes may also appear in morphologically non-transparent words, as shown in (365). In such cases, there is a single word stress, which may fall on the prefix (365a). Vowel reduction may apply to the vowel corresponding to the stressed vowel in the transparent counterpart (365b versus 364a). The lowering of the stressless vowel before the prefix final consonant may no longer apply (365c versus 364b), and the initial diphthong of *extra-* becomes impossible in an unmarked pronunciation (365d versus 364c).

- (365) a. *polÍgamo* ‘polygamist’
 hipÉrbole ‘hyperbole’
 supÉrfluo ‘superfluous’
 b. *poliedro* [u] ‘polyhedron’
 c. *internacional* [i] ‘international’
 d. *extraordinário* [ʃ] ‘extraordinary’

What is crucial to point out is that in transparent formations, and unlike in the non-transparent cases, the properties that indicate the presence of a word stress always characterize these prefixes.

As for the presence of an independent word-stress on the base, the same tests already referred to for prefixed words with monosyllabic prefixes show that the morphological base is also assigned a word-stress. The inter-

nal prosodic structure of these derived words should thus be represented as in (366).

- (366) (inter)_ω (cultural)_ω
 (mega)_ω (concerto)_ω

In the light of (362) and (366), the next question we must deal with is the identification of the prosodic node that dominates both members of these derived words. As before, two possibilities have to be evaluated: either the two prosodic words that form the prefixed word are grouped together under a ϕ -phrase, or they form a compound prosodic word and are thus dominated by a ω^{\max} .

There are two major arguments favoring the compound prosodic word analysis. The first one follows from the behavior of these prefixed words with respect to final central vowel deletion (see chapter 3, section 14). Similar to round vowel deletion and non-back vowel deletion, central vowel deletion is obligatorily blocked if the second vowel bears word stress and belongs to the head of the compound. This is shown in (367).

- | | | | | |
|-------|-----------------------------|--------|--------------------------------|-------|
| (367) | <i>extra-época</i> | *0/[ɐ] | <i>extra-eleições</i> | 0/[ɐ] |
| | ‘out of the regular period’ | | ‘out of the elections context’ | |
| | <i>ultra-óbvio</i> | *0/[ɐ] | <i>ultra-ocupado</i> | 0/[ɐ] |
| | ‘extremely obvious’ | | ‘extremely busy’ | |

Crucially, this behavior obtains regardless of ϕ -level prominence, as illustrated in (368a) below, where the second prosodic word of the prefixed word is not the head of the phonological phrase. Example (368a) contrasts with (368b), since in the latter case vowel deletion is not totally excluded. In (368b) the two relevant prosodic words do not form a compound prosodic word and neither of them constitutes the head of the phonological phrase, which has been described to correspond to the rightmost prosodic word in the ϕ -domain (e.g. Frota 2000, and chapter 1, section 2 of this book).²⁹¹

- (368) a. (um ultra-Óbvio problema)_φ *levantado por alguém* *0/[ɐ]
 ‘an extremely obvious problem raised by someone’
 b. (a quinta Ordem dada)_φ *não chegou a ser obedecida* ?0/[ɐ]
 ‘the fifth order given was never obeyed’

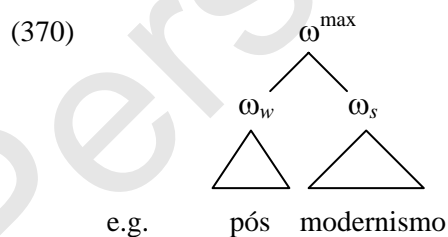
232 *Compound prosodic words*

The second argument for the compound prosodic word approach is provided by focus assignment. As shown in (369), focal stress may only be assigned to the second prosodic word of the construction.

- (369) a. A: *O teu irmão comprou uma vassoura ontem, não foi?*
 ‘Your brother buy a sweeper yesterday, didn’t he?’
 B: *Não. O meu irmão comprou um mini-aspira**DOR** ontem*
 (*não uma vassoura*).
 B’: **Não. O meu irmão comprou um **MI**ni-aspirador ontem*
 (*não uma vassoura*).
 ‘No. My brother bought a mini-vacuum-cleaner yesterday
 (not a sweeper).’
- b. A: *O Almada foi o fundador do surrealismo em Portugal.*
 ‘Almada was the initiator of surrealism in Portugal.’
 B: *Não concordo nada. Eu acho que foi o fundador do*
*pós-modern**IS**mo em Portugal* (*não do surrealismo*).
 B’: **Não concordo nada. Ele foi o fundador do*
***PÓS**-modernismo em Portugal* (*não do surrealismo*).
 ‘I completely disagree. He was the initiator of
 post-modernism in Portugal (not surrealism).’

Our interpretation of these facts is similar to the interpretation given in the preceding section for parallel cases with suffixed words. If we assume that the members of these derived words form together a compound prosodic word whose head corresponds to the rightmost prosodic word, focal stress association with the second member of the construction follows straightforwardly from its status as the head of the compound prosodic word.²⁹²

To sum up, according to the diagnostics for prosodic wordhood surveyed in this section, derived words with stressed prefixes consist of two prosodic words, which form together a compound prosodic word, as depicted in (370).



3.3. Morphological compounds (*stem compounds*)

In this section we investigate the prosodization of compound words that have a stem as the first member of the compound. Following the spirit of Villalva's (1994) terminology, we will call these structures *morphological compounds*.

We have seen in chapter 1, section 6.4 that a distinction may be drawn within this type of compounds between *stem+stem* compounds and *stem+word* compounds. This distinction has been argued to have consequences for the prosodization of the two types of words in different languages (e.g. Ralli and Nespor 1996 and Nespor 1999b). However, in her study of the morphological structure of Portuguese words, Villalva (1994) explicitly rules out the possibility of stem+word composition in EP. In the description that follows, which is carried out from a strictly phonological point of view, we will consider any compound whose first member is a stem. The relation between the morphological structure of these words and their phonological behavior will be discussed later on in this section.

As illustrated in (371) and (372) below, some morphological compounds appear to constitute single stress domains, while others seem to form two stress domains. The presence of a single word stress in the forms in (371) is evident: only one word stress is perceived; the main stress may fall on the linking vowel of the compound or on the final vowel of the first stem (see examples 371b–d)²⁹³; the vowel of the first stem, which has a stressed counterpart in the compounds of the type shown in (372), may undergo vowel reduction (examples 371a and d) and semivocalization (371c), and it may show variable realizations when word initial (371e).²⁹⁴

(371) a.	<i>fotografia</i>	[futugrɐfɨɐ]	'photograph'
b.	<i>autónomo</i>	[awtɔnumu]	'autonomous'
c.	<i>biólogo</i>	[biɔʎlugu]/[bjɔʎlugu]	'biologist'
d.	<i>telégrafo</i>	[tɨlɛgrɐfu]	'telegraph'
e.	<i>helicóptero</i>	[elikɔptiru]/[ilikɔptiru]	'helicopter'

The presence of two word stresses in the forms in (372) is also clear: the two stresses are perceived as such; the stressed vowel of each stem cannot undergo vowel reduction (e.g. 372a and d) or semivocalization (see 372c for the first stem); and it does not show variation in height when it occurs word initially (see 372f for the first stem). Besides word stress, two other properties contribute to the delimitation of the first stem in the forms in

(372). The final vowel of the stem (or the linking vowel) systematically fails to undergo vowel reduction (see also chapter 3, section 3), unlike in compounds that form a single stress domain (see 372a–d versus 371a–b). In addition, the final unstressed vowel of the first stem is lowered in syllables closed by /r/, as in (372e). Given that this process applies at the right-edge of the prosodic word domain, this shows that the first stem is an independent prosodic word (see chapter 3, section 8).

(372) a. <i>foto-montagem</i>	[fótɔmõtázẽ]	‘photocomposition’
b. <i>auto-admiração</i>	[áwtɔɐdmirɛsẽw]	‘self-admiration’
c. <i>bio-médico</i>	[bíɔmédiku]	‘biomedical’
d. <i>téle-conferência</i>	[télekõfĩrẽsjɐ]	lit. ‘distant conference’
e. <i>infor-jovem</i>	[ĩfɔrʒóvẽ]	lit. ‘information youth’
f. <i>héli-porto</i>	[élipórtu]	‘heliport’

At this point, we may address the issue introduced above concerning the distinction between stem+stem compounds and stem+word compounds.

If it is assumed that the prosodic difference illustrated in (371) and (372) is the result of two distinct morphological structures associated with the compounds at stake, then our observations appear to support a morphological distinction among morphological compounds: (i) in the forms in (371) the second member of the compound does not correspond to existing words, and thus these compounds may be formed by the concatenation of two stems; (ii) in the items in (372), by contrast, the second member of the compound corresponds to an existing word, and thus these compounds may be formed by the concatenation of a stem plus a word. Under an analysis that assumes a morphological distinction between types of morphological compounds, there would be a straightforward correlation between the morphological and the phonological data, as proposed in Nespor and Ralli (1996) and Nespor (1999b) for other languages.

There are however at least two arguments against this analysis in EP. First, words like *poligâmico* ‘polygamist’ include two prosodic words ((*poli*)_ω (*gâmico*)_ω) although *gâmico* is not an existing word. Therefore, here the prosodic word status of the second part of the compound is not correlated with its morphological word status. Second, according to Villalva (1994) there are no stem+word compounds in EP. The argument follows from the observation that words such as *macro-economista* ‘macro-economist’ may not involve the concatenation of the stem *macro* with the word *economista* given the meaning of the whole expression: *macro-*

economista means ‘a person who deals with macro-economy’ rather than ‘a big economist’. Villalva (1994: 355) thus proposes for this word the morphological structure in (373a), instead of (373b).

- (373) a. [[[[macro] [econom]] [ist]] a]
 b. *[[macro] [[econom] [ist]] a]

Thus, assuming Villalva’s analysis of the morphological structure associated to these compounds, there is a mismatch between the morphological and the prosodic structure, since a unit like *economista* forms a prosodic but not a morphological constituent. We will return to this issue in the final section of this chapter, where we discuss the mapping between morpho-syntactic and prosodic structures.

With respect to the first group of morphological compounds (illustrated in 371 above), the phonological criteria point to the prosodic grouping of the two stems under the same prosodic word, as represented in (374).

- (374) (fotografia)_ω
 (biólogo)_ω

As for the second group of morphological compounds (illustrated in 372), each member of the compound behaves like an independent prosodic word (in the case of the second member, the prosodic word includes the stem and any following suffix). This is represented in (375).

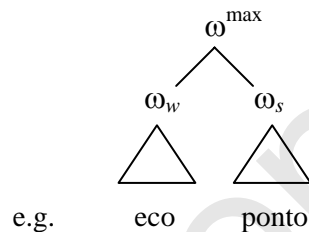
- (375) (foto)_ω (composição)_ω
 (bio)_ω (médico)_ω

The issue we must address now concerns the identification of the prosodic node that dominates the two prosodic words in the latter cases. Due to the limited data available, we can only use one test in order to support the compound prosodic word analysis, which is focus assignment.²⁹⁵ Like in the constructions considered so far, focus may only be assigned to the second prosodic word of the morphological compound, as shown by the examples in (376).

- (376) a. A: *Disseram-me que o João é o responsável pela publicidade desta página.*
 ‘I was told that John is responsible for the publicity of this page.’
 B: *Mas não: O João é o responsável pela foto-mon**TA**gem desta página.*
 B’: **Mas não: O João é o responsável pela **F**Oto-montagem desta página.*
 ‘But he isn’t: John is responsible for the fotocomposition of this page.’
- b. A: *O Zé despejou as garrafas num contentor repleto de dejectos.*
 ‘Zé put the bottles in a wagon full of trash.’
 B: *Nada disso. O Zé despejou as garrafas num eco-**PON**to repleto de dejectos.*
 B’: **Nada disso. O Zé despejou as garrafas num **Eco**-ponto repleto de dejectos.*
 ‘No he didn’t. Zé put the bottles in a recycle-bin full of trash.’

We take this as evidence for the prosodization of these words as compound prosodic words, like in (377).

(377)



3.4. Syntactic compounds (word compounds)

The morphosyntactic structure associated to compounds formed by the concatenation of two (or more) words is also a matter of some controversy. Along the lines of Di Sciullo and Williams (1987), Nespor and Ralli (1996) assume that word+word compounds result from the reanalysis of syntactic phrases as X^0 units. Villalva (1994), by contrast, proposes that this type of reanalyzed structures in EP only characterizes sequences formed by a verb and a complement or by two verbs. The latter type of composition is con-

sidered to be morphosyntactic, since the structure of the base is a syntactic constituent reanalyzed as a stem. The remaining cases of word compounds are argued by Villalva to form either right-adjunction structures (consisting of two nouns, where the first member is the nucleus of the compound) or conjunction structures (formed by the coordination of two or more nouns or adjectives). Certain expressions are seen, in addition, to be lexicalized syntactic expressions (like those formed by a noun and a prepositional phrase, or by a noun and an adjective). Peperkamp (1997a), in turn, proposes that in Italian the only type of productive composition is verb+noun compounding. As for left-headed noun+noun compounds, they are assumed to form listed phrases, whereas the remaining types of compounds are considered to be lexicalized, resulting from processes that are no longer productive.

In the phonological description that follows, we will consider compounds formed by two words, regardless of their morphosyntactic structure. We will see, nevertheless, that such compounds do not appear to behave uniformly from a phonological point of view. As before, our description is restricted to the available phonological data.

Like in the preceding case, some words that are historically composed of two words may be phonologically treated like a single prosodic word (e.g. *pernalta* [pɛrnáʎtə] ‘a bird of the order Gallator’ < *perna* [pérnə] ‘leg’, *alta* [áʎtə] tall-FEM). This type of word may be seen to have been reanalyzed as a single word through an unpredictable process of lexicalization. In general, however, word+word compounds are composed of two or more independent prosodic words. We will concentrate our investigation on the latter cases.

The demonstration that each member of word+word compounds may form an independent prosodic word seems redundant, given the uncontroversial nature of this statement – as mentioned throughout chapter 2, the presence of two word stresses in the description of these constructions is repeatedly found in EP literature (e.g. Gonçalves Viana 1883, Morais Barbosa 1965, d’Andrade 1994b, Pereira 1999). In fact, in the cases under observation the words that form a compound have the same phonological shape as when they appear in isolation. We will thus take for granted the prosodic word status of each member of the word-compounds considered in the following paragraphs.

The focus of our description will hence be the identification of the prosodic constituent that dominates the two prosodic words corresponding to a word-compound. As before, two possibilities will be evaluated: either

238 *Compound prosodic words*

the prosodic words are immediately dominated by the phonological phrase node, or by a prosodic node of the same level, in a recursive compound word structure.

Let us first consider the facts of vowel deletion. As illustrated in (378), all vowel deletion processes are blocked if the target vowel is followed by a vowel that bears word stress (378a), while they regularly apply if V2 does not bear word stress (378b).²⁹⁶

- | | | | | |
|----------|--|--|---|--|
| (378) a. | i. <i>sAlto Alto</i> [w]/*0
'high heel' | ii. <i>pOrta-Óculos</i> [ɐ]/*0
'glasses holder' | iii. <i>sANto hOmem</i> [w]/*0
'holy man' | iv. <i>grANde Área</i> [j]/*0
'penalty area' |
| b. | i. <i>rIso amarElo</i> [w]/0
'yellow smile' | ii. <i>pOrta-objEcto</i> [ɐ]/0
'object holder' | iii. <i>pequEno-almOço</i> [w]/0
lit. 'little lunch' | iv. <i>grANde-oficiAl</i> ^x [j]/0
lit. 'big officer' |

At least in two of these examples, namely (378a.ii) and (378a.iv), the (im)possibility of vowel deletion if the second member of the compound is not in ϕ final position is very clear: vowel deletion is obligatorily blocked regardless of ϕ -level prominence, as illustrated in (379).

- (379) *O João comprou (um porta-óculos castanho) ϕ* [ɐ]/*0
'John bought a brown glasses holder'
- Vimos (a grande área contrária) ϕ cheia de jogadores* [j]/*0
'(We) saw the opposite penalty area full of players'

By contrast, in the examples in (380) the obligatory blocking of vowel deletion if the second member of the compound does not correspond to the head of a ϕ is not that clear.²⁹⁷

- (380) *A moça trazia uns sapatos (com um salto alto elegante) ϕ* (?) [w]/0
'the girl was wearing shoes with elegant high heel'
- (Um santo homem cansado) ϕ também peca* (?) [w]/0
'a tired holy man also offends'

With respect to non-back vowel deletion, some compounds behave differently from *grande área*, as shown in (381): if V2 bears word stress,

vowel deletion is not obligatorily blocked (381a); in addition, if the compound is followed by a word within the same phonological phrase, [j] no longer yields a natural realization (381c).²⁹⁸

- (381) a. *vErde Água* ([j])/0 lit. ‘green water’
 b. *vErde azeitOna* ^x[j]/0 lit. ‘green olive’
 c. *Tinha olhos* (de um *vErde Água* translúcido)_φ ^x[j]/0
 ‘(He) had the eyes of a translucent green like the water’

The facts noted above point to the prosodization of the words in (378a.ii) and (378a.iv) as compound prosodic words. The words in (378a.i–iii) and (381a), by contrast, do not clearly behave as prosodic compounds, and therefore each member of the word may be directly dominated by the ϕ node.

The prosodic facts described correlate with the morphosyntactic structure of compounds in the following way. The forms that behave like prosodic compounds are V+N and A+N compounds, whereas the forms that do not clearly behave like prosodic compounds are N+A, A+N, and N+N expressions. The fact that A+N compounds may behave phonologically as prosodic compounds or as sequences of prosodic words dominated by the ϕ -node suggests that the explanation for the prosodization of *grande área* (‘penalty area’ and not ‘big area’) as a prosodic compound may have to do with an unpredictable process of lexicalization of this construction, which caused its meaning and morphological structure to become opaque. An argument in favor of this view is the behavior of this compound with respect to inflection: while the plural form of other right-headed A+N compounds, such as *santo homem* and *pequeno almoço* (see 378 above), is *santos homens* and *pequenos almoços* (i.e. each member of the compound is inflected), the plural of *grande área* is *grande áreas* [grẽdʒárjɐs] (i.e. the first member of the compound does not inflect, for the speakers that treat this as a non-transparent compound).

Unlike in the case of A+N compounds, we found no counter-example to the prosodization of V+N formations as compound prosodic words. We will put forward the hypothesis that this type of compound regularly forms a compound prosodic word. The contrast between this type of composition and the remaining cases that may also form an X⁰-level syntactic unit may follow from the fact that these are the only syntactic compounds (together with V+V compounds, which are rather rare in EP) whose internal syntactic structure is both morphologically and syntactically opaque (cf. Villalva

1994: 6.2.2). This is shown by the absence of morphological inflection of the nucleus of the syntactic expression that gives rise to the compound (i.e. the verb) and the impossibility of morphological or syntactic modification of the members of the compound.

The test of focus assignment confirms our observations concerning V+N compounds. As illustrated in (382), focal stress may only be assigned to the rightmost prosodic word of the syntactic compound, which we propose to be the head of the phonological compound.²⁹⁹

(382) A: *O João ofereceu um estojo à Maria.*

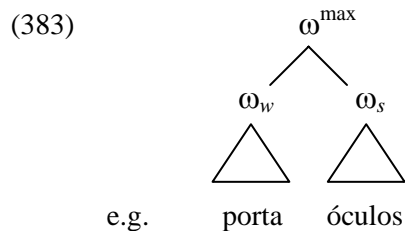
‘John offered a case to Mary.’

B: *Não senhor. O João ofereceu um porta-Óculos à Maria.*

B’: **Não senhor. O João ofereceu um PORTa-óculos à Maria.*

‘Not at all. John gave a glasses holder to Mary.’

The proposed prosodic structure associated to V+N compounds is thus the one given in (383).



Before we conclude, we should observe that the data surveyed in this section is limited and more investigation is required on the different possibilities of prosodization available for each type of word+word compound, as well as on the exact relation between morphosyntactic and prosodic structure.

To sum up, besides V+N compounds, which appear to regularly form a compound prosodic word, other combinations of words may be similarly prosodized, possibly as a consequence of a lexicalization process (that is, if the syntactic structure of the whole expression becomes morphologically and syntactically opaque).

3.5. Abbreviations

In addition to the constructions investigated so far, most abbreviations show a phonological behavior distinct both from words that form a single word stress domain and from prosodic words that are directly dominated by the phonological phrase node.

As before, there are a few abbreviations that have been lexicalized as a single prosodic word (e.g. *laser* 'id.', *ovni* 'UFO'). In these cases, speakers do not analyze the word as an abbreviation. In most cases, however, the name of each of the letters that form the abbreviation corresponds to a prosodic word. Besides the perception of stress on each word, the fact that vowel reduction and semivocalization never affect the stressed vowels of these words demonstrates their prosodic word status, as illustrated in (384).

- (384) *PS* [péés]/*[pjés] (abbreviation of a political party name)
BD [béde]/*[bidé] (abbreviation of *banda desenhada* 'cartoon')
PT [pété]/*[pité] (abbreviation of a company name)
CD [séde]/*[sidé] (abbreviation of 'compact disk')

Abbreviations provide a particularly suitable case for the testing of the compound prosodic word analysis: (i) it is possible to verify the application/blocking of non-back vowel deletion and central vowel deletion (as combinations of letter names that end in non-back or central vowels with letter names that start with a stressed vowel are common); (ii) there are many examples of abbreviations consisting of more than two potential prosodic words, thus allowing us to determine the exact relevance of the head of the prosodic compound for the blocking of these processes; (iii) the judgments on the data are very clear; and (iv) the behavior of abbreviations is extremely regular.

We will first consider non-back vowel deletion. As the examples in (385) show, non-back vowel deletion in prosodic word final position is always obligatorily blocked if the vowel that follows bears the rightmost word stress of the abbreviation.³⁰⁰

- (385) *RN* (*erre ene*) [j]/*0 ([éRjén])
 (abbreviation of a company name)
UMM (*u eme eme*) [j]/*0 ([úémjém])
 (brand of a jeep)

242 *Compound prosodic words*

TSF (tê esse efe) [j]/*0 ([téésjéf])
 (abbreviation of a radio station name)

That this does not follow from ϕ -level prominence is shown in (386), where vowel deletion is still impossible although the abbreviation is followed by a prosodic word within the same phonological phrase.

(386) *Fala-se sobre a criação* (de uma TSF regional) _{ϕ} (tê esse efe) [j]/*0
 ‘they talk about the creation of a local TSF’

This type of behavior was already found in derived words with two stress domains (section 3.1), and in some word compounds (section 3.4). Like before, our interpretation of these facts is that non-back vowel deletion is obligatorily blocked within a compound prosodic word when the target vowel is followed by a vowel bearing the compound prosodic word stress.

The relevance of the head of the compound prosodic word can be further seen in the examples in (387), which consist of three prosodic words. While vowel deletion tends to apply between the first prosodic word and the second prosodic word, even if V2 bears word stress, it is obligatorily blocked between the second and the third prosodic words when V2 is stressed. In our analysis, this is so because only in the latter case does V2 bear the compound prosodic word stress.

(387) *RFM* (erre efe eme) [ÉRÉfjém]/*[ÉRÉfém]
 (name of a radio station)
MFA (eme efe a) [éméfjáj]/*[éméfá]
 (abbreviation of the name of a political party)

That the domain of non-back vowel deletion is the compound prosodic word is also shown by the regular deletion of non-back vowels that are final within the compound. This may be seen in the examples above that show a non-back vowel in this position, namely, *RFM* and all the examples in (385).

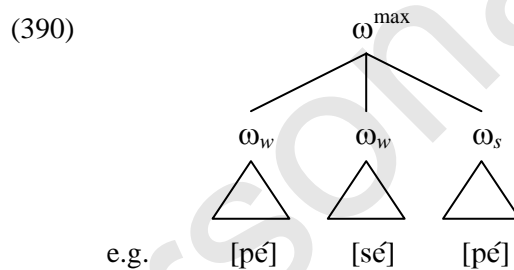
A similar picture obtains with final central vowel deletion (see 388). If V2 is stressed and belongs to the rightmost prosodic word of the compound, the preceding central vowel is not deleted (see 388a), regardless of ϕ -level prominence. By contrast, deletion becomes optional if V2 is stressed and belongs to a compound internal prosodic word (388b).

- (388) a. *JL* (*jota ele*) [ɐ]/*0 ([ʒótɐɛ́t̚]/*[ʒótɛ́t̚])
 (abbreviation of a newspaper name)
 b. *JSD* (*jota esse dê*) [ɐ]/0 ([ʒótɐɛ́sdɛ́]/[ʒótɛ́sdɛ́])
 (abbreviation of a political party name)
KLM (*capa ele eme*) [ɐ]/0 ([káɸɛ́ljém]/[káɸɛ́ljém])
 (name of an airline company)

The test of focus assignment further corroborates this analysis. As illustrated in (389), within an abbreviation formed by three prosodic words, focal stress may only be assigned to the rightmost prosodic word of the compound.

- (389) A: *Dizem que o João era entusiasta dos ecologistas quando era jovem.*
 ‘(They) say that John was an enthusiast of the ecologist party when (he) was young.’
 B: *Nada disso. Ele era entusiasta do PCP (*PÊ cê pê / *pê CÊ pê / pê cê PÊ) quando era jovem.*
 ‘Not at all. He was enthusiast of PCP (a political party) when he was young.’

The proposed prosodic structure associated to abbreviations composed of two or more prosodic words is thus a compound prosodic word structure, as illustrated in (390) for the abbreviation *PCP*.³⁰¹



3.6. Mesoclititic structures

As we have seen in chapters 2 and 4, the construction traditionally considered to yield mesoclititic structures, that is structures with *internal* pronomi-

nal clitics, also displays two word stresses. Examples of such formations are provided in (391). Besides perception, the presence of two word stresses is demonstrated by the obligatory absence of vowel reduction in the infinitival form (391a–c) and in the affix-like element *-haver* (391a–b).

- | | |
|-----------------------------|------------------------------|
| (391) a. <i>falAr-te-Ei</i> | c. <i>percebÊ-lo-Íamos</i> |
| ‘(I) will speak to you’ | ‘(we) would understand him’ |
| b. <i>comÊ-la-Emos</i> | d. <i>servIr-ma-Ías</i> |
| ‘(we) will eat it’ | ‘(you) would serve it to me’ |

The presence of two prosodic words in this construction is therefore unproblematic. By contrast, the organization of these prosodic words, as well as the prosodization of the internal clitic require detailed investigation.

We will set out to determine the direction of cliticization of the pronominal clitic in this construction. As we will see in the following discussion, clitic final vowels behave like prosodic word final within a compound prosodic word, rather than like vowels belonging to proclitic function words.

The contrast between the behavior of these pronouns and the behavior of proclitic function words can be seen in the examples in (392). While the non-back vowel of the clitic pronoun may never be deleted in mesoclitic structures when followed by a vowel (392a), there is always [j]/0 alternation in similar environments involving a proclitic function word and its phonological host (392b).

- | | | |
|-------------------------------|--------|----------------------------------|
| (392) a. <i>dAr-me-Ás</i> | [j]/*0 | ‘(you) will give (it) to me’ |
| <i>pensAr-se-Á</i> | [j]/*0 | ‘(it) will be thought’ |
| <i>cantAr-te-Emos</i> | [j]/*0 | ‘(we) will sing to you’ |
| b. <i>vais gostAr de Aves</i> | [j]/0 | ‘(you) will like birds’ |
| <i>jÁ se Abre a porta</i> | [j]/0 | ‘the door can already be opened’ |
| <i>vão pensAr que Eras tu</i> | [j]/0 | ‘(they) will think (it) was you’ |

Notice that in a few particular cases involving proclitic function words, the alternation [j]/0 is no longer possible due to the lexicalization of the expression that comprises the clitic (see also chapter 7, section 5.4.2). As illustrated in (393), in these cases the realization that becomes obligatory is the reduced one, in contrast with the obligatory realization of the glide in mesoclitic constructions.

- (393) *deste* (<de este) [dɛʃt] lit. 'of this'
mãe d'água (<de água) [mɐ̃j'dágwɐ] lit. 'mother of water'

The data above suggest that the clitic in mesoclitic constructions is not proclitic. The facts of final vowel deletion show, furthermore, that the clitic behaves like an element that belongs to the preceding prosodic word, within a compound prosodic word domain. Indeed, the obligatory presence of the glide in the clitic forms can be seen to follow from the obligatory blocking of non-back vowel deletion when the target vowel is followed by a prosodic word starting with a vowel that bears the prominence of a compound prosodic word. Since the second prosodic word of the mesoclitic formation always starts with a stressed vowel, which belongs to the rightmost prosodic word of the whole construction, and given the assumption that these clitics – like other postverbal pronominal clitics – are incorporated into the prosodic word that contains the verb, non-back vowel deletion is predicted to always be blocked in these cases. As we have seen, the prediction that emerges from the compound prosodic word analysis is borne out by the data. The proposed prosodic structure of mesoclitic formations is given in (394).³⁰²

- (394) ((cantar-te)_ω (emos)_ω)_ω^{max}

Notice that, as before, ϕ -level prominence does not affect the blocking of non-back vowel deletion. This is shown in (395), where the mesoclitic construction is followed by another word within the same ϕ .

- (395) *os alunos* (oferecer-lhe-ão flores)_φ [j]/*0
 '(they) will offer you flowers'
os teus amigos (emprestar-te-ão livros)_φ [j]/*0
 'your friends will lend you books'

The behavior of these clitics with respect to the phonological process of round vowel deletion further supports the compound prosodic word analysis. Although round vowel deletion may affect enclitic pronominals in non-compound prosodic configurations, as in (396a), it is obligatorily blocked if V2 bears the prosodic compound's prominence (see 396b).

- (396) a. *estava a achá-lo enervado* [w]/0 '(I) was thinking he was nervous'

246 *Compound prosodic words*

- b. *percebê-lo-emos* [w]/*0 ‘(we) will understand it’
escrevê-lo-ia [w]/*0 ‘(he) would write it’

Again, in mesoclititic constructions the rule is blocked regardless of the locus of ϕ -level prominence, as illustrated in (397).

- (397) *se pudesse, o João (emprestá-lo-ia hoje) ϕ à Maria* [w]/*0
 ‘if (he) could, John would lend it today to Mary’
(di-lo-emos todos) ϕ quando for possível [w]/*0
 ‘(we) will all tell it, when (it) will be possible’

Central vowel deletion is another process blocked in similar environments, as shown in (398).

- (398) *a gata, (deixá-la-ia aqui) ϕ* [ɐ]/*0
 ‘the she-cat, (he) would leave it here’
a revista, (lê-la-emos logo) ϕ [ɐ]/*0
 ‘the magazine, (we) will read it later’

Notice that this process shows in addition that the whole compound behaves like any prosodic word, since it optionally applies to a compound final vowel, when followed by a word also starting with a vowel, as in (399).³⁰³

- (399) *achá-la-ia engraçada* [ɐ]/0 ‘(I) would find her amusing’
dir-te-ia obrigado [ɐ]/0 ‘(I) would say thanks to you’

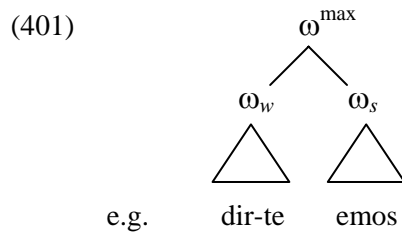
Finally, focus distribution also patterns as expected if the compound prosodic word analysis is assumed. As illustrated in (400), focal stress may only be assigned to the rightmost prosodic word of the prosodic compound, that is to the head of this constituent.

- (400) A: *A Maria dirá ao João que aqui estamos.*
 ‘Mary will tell John we are here’
 B: *Não senhor. Quando muito repeti-lo-Á ao João, porque ela já o disse.*

B': **Não senhor. Quando muito repeTI-lo-á ao João, porque ela já o disse.*

'No she won't. At best, she will repeat it, because she has already told it to him'

The data presented in this section provides clear support for the prosodization of mesoclitic constructions as compound prosodic words. As for the internal clitic, it patterns like other postverbal pronominal clitics, being incorporated into the preceding prosodic word. This is depicted in (401).



3.7. Specific combinations of words

In addition to the constructions considered so far, which regularly behave like compound prosodic words, there are also some specific combinations of words that display a similar phonological pattern. This is indicated in particular by the facts of non-back vowel deletion: similar to some derived and compound words, to abbreviations and to mesoclitic structures, in these cases too non-back vowel deletion is obligatorily blocked when V2 is stressed and final within the expression. There are two classes of cases to be considered: (i) those in which the elements that obligatorily display vowel deletion blocking are totally predictable; (ii) and those in which the elements that obligatorily display vowel deletion blocking are not entirely predictable.

Let us consider each of the two cases in turn, starting with the type mentioned in (i) above, and illustrated in (402). As shown by the examples in (402a), which include the typical combinations of numbers and letters found in Portuguese car license plates, the pair of letter names always displays non-back vowel deletion blocking. The same pattern obtains in the names of roads composed of a letter name (N for *nacional* 'national') and a number, as shown in (402b).

248 *Compound prosodic words*

- (402) a. 54-25-LM (*ele Eme*) [j]/*0 b. N1 (*ene UM*) [j]/*0
 73-92-ML (*eme Ele*) [j]/*0 N8 (*ene Oito*) [j]/*0
 43-56-FR (*efe Erre*) [j]/*0 N11 (*ene ONze*) [j]/*0

The relevance of the stress in V2 for the blocking of the rule is confirmed by the application of non-back vowel deletion when V2 is unstressed, as in (403).

- (403) 67-23-FH (*efe agÁ*) ^x[j]/0
 56-98-LH (*ele agÁ*) ^x[j]/0

In order to account for these data, we propose that these cases also display compound prosodic word structures. The generalization we can establish is that a compound prosodic word obtains when the first member of the compound is the name of a letter, while the second member may be either the name of a letter or a numeral.

Numerals are also involved in the second class of cases mentioned above. Here the observed pattern depends on the specific words that are combined, as shown by the examples in (404). The numerals in (404a–a') display compound prosodic word behavior, since non-back vowel deletion is blocked when V2 bears the compound stress. In fact, as illustrated in (404d), if V2 is not stressed, the vowel is regularly deleted. However, the compound pattern only arises if the second word is 'horas' or 'anos', as shown by the contrast between the forms in (404a–a') and those in (404b).³⁰⁴ In addition, the same behavior is not found with some other numerals ending with a non-back vowel, as those in (404c).

- (404) a. *onze hOras* *0/[j] 'eleven hours'
 doze hOras *0/[j] 'twelve hours'
 vinte hOras *0/[j] 'twenty hours'
 a'. *onze Anos* *0/[j] 'eleven years'
 doze Anos *0/[j] 'twelve years'
 vinte Anos *0/[j] 'twenty years'
 b. *onze Ursos* 0/^x[j] 'eleven bears'
 doze Aves 0/^x[j] 'twelve birds'
 vinte Obras 0/^x[j] 'twenty works'
 c. *sete hOras* 0/^x[j] 'seven hours'
 nove hOras 0/^x[j] 'nine hours'

- d. *onze horItas* 0/^x[j] eleven hours-DIM
doze anInhos 0/^x[j] twelve years-DIM

There is at least one difference between the numerals in (404), which may be correlated with their different prosodizations. The numerals that are prosodized within a compound prosodic word together with *horas* and *anos* form the closed class (with just six members: *onze* ‘eleven’, *doze* ‘twelve’, *treze* ‘thirteen’, *catorze* ‘fourteen’, *quinze* ‘fifteenth’, and *vinte* ‘twenty’) of numerals that end in a non-back vowel and that do not occur as the second member in the formation of other numerals (unlike *sete* and *nove*, e.g. *trinta e sete* ‘thirty seven’, *trinta e nove* ‘thirty nine’, *quarenta e sete* ‘forty seven’, *quarenta e nove* ‘forty nine’, and so on). As for the specificity of the words *horas* and *anos*, two factors should account for the fact that they establish a compound relation with some numerals: first, they typically occur with words of this class (in addition to demonstratives and prepositions); second, they are highly frequent words in EP. In fact, together with *hoje* ‘today’ (which may not appear with numerals), these are the most frequent content words that start with a stressed vowel, according to the list of EP most frequent words in Bacelar, Marques, and Segura da Cruz (1987). Their ranking (where 1=the most frequent word) and number of occurrences in the corpus collected (containing 700 000 words) are shown in (405).

(405)

	Ranking	Frequency
<i>anos</i>	42	2124
<i>horas</i>	114	642

All these facts point to the analysis of these expressions as forming together a compound prosodic word because of a lexicalization process (as in the case of syntactic compounds). That is, here we assume that the specific prosodization obtained does not follow directly from the regular syntactic structure displayed by these sequences of words.

To sum up, in this section we have seen that some specific sequences of prosodic words may be grouped together within a prosodic compound: in particular, sequences of prosodic words consisting of pairs of letter names (like those found in car license plates), pairs of prosodic words consisting of names of letters followed by numerals, as well as (some) numerals followed by the very frequent words *horas* and *anos*.

3.8. *Interim summary*

In section 3 we have investigated the phonological behavior displayed by a number of morphosyntactic constructions that include more than a single stress unit. These constructions comprise (i) derived words with suffixes that form stress domains independent of their base, (ii) derived words with stressed prefixes, (iii) morphological (stem plus stem) compounds, (iv) morphosyntactic and syntactic (word plus word) compounds, (v) abbreviations, and (vi) mesoclititic structures. Sequences of prosodic words that display similar phonological behavior, consisting of (a) pairs of letter names, (b) names of letters followed by numerals, as well as (c) (some) numerals followed by the words *horas* and *anos*, have also been observed. In all these cases, we found sound evidence for the prosodic word status of the internal members involved in each construction. To the exception of syntactic compounds, we also offered clear-cut evidence for the prosodic grouping of the internal prosodic words of these constructions into a compound prosodic word. The prosodization of the sequences of prosodic words involved in these constructions is thus seen to explain their peculiar behavior when compared with other sequences of prosodic words immediately dominated by the phonological phrase node.

Within the class of word compounds, phonological evidence has led us to propose that only morphosyntactic (i.e. V+N) composition regularly displays phonological composition. Other word plus word combinations, by contrast, were hypothesized not to display obligatory phonological composition, unless a lexicalization process has caused the syntactic and morphological structure of the compound to become opaque.

4. Additional evidence for the prosodic word domain

In the present section we add two final pieces of evidence for the prosodic word status of the internal members of prosodic compounds, as well as for the prosodic word domain in general. Deletion under identity and clipping provide such evidence. Besides the fact that these phenomena constitute additional potential tests for the prosodic word domain, they also allow us to highlight the contrast between constructions that are morphosyntactically alike but prosodically distinct. We should remark, however, that at this point we do not wish to present a full account of the mechanisms involved in deletion under identity or clipping, which may not be purely phonological. Also, we will only concentrate on the phonological issues.

Let us start by looking at deletion under identity (cf. Booij 1985, 1995, Wiese 1993, 1996, Kleinhenz 1994, and chapter 1, section 6.1 for an illustration of the process in Dutch). The interest of this process is twofold: (i) on the one hand, it will be shown that the unit to be deleted appears not to be defined strictly in morphosyntactic terms, but rather that it systematically corresponds to a prosodic constituent – the prosodic word domain; (ii) on the other hand, the observation of this process will reveal prosodic differences between similar morphosyntactic structures.

Essentially, deletion under identity (DUI) – or coordination reduction – consists of the deletion of a unit within a complex word in a coordinate structure with partial phonological identity (see the examples in 406a below). The relevance of the prosodic word status of the target elements in the process of DUI is shown by the contrast between the examples in (406a) and the examples in (406b). Both groups of words consist of derived words within coordinated structures, where the first and the second members of the coordinate structure are partially similar.

- (406) a. ~~seguramente~~ mas lentamente > *segura mas lentamente*
 ‘surely but slowly’
 pré-~~guerra~~ e pós-guerra > *pré e pós-guerra*
 ‘before war and after war’
 mono-~~acentual~~ ou bi-acentual > *mono ou bi-acentual*
 ‘with one or two stresses’
 b. ~~acampamento~~ e acantonamento > **acampa e acantonamento*
 ‘camping and sheltering’
 ele repensou e ~~reescreveu~~ > **ele repensou e escreveu*
 ‘he thought again and wrote again’
 ele des~~fez~~ e refez > **ele des e refez*
 ‘he has undone and redone’
 em pró~~clise~~ ou ênclise > **em pró ou ênclise*
 ‘in proclisis or enclisis’

Nevertheless, DUI is only possible in the examples in (406a). Given the similarity of the words involved from a morphological point of view, the explanation for this contrast appears not to be found in the morphological structure of the words. The words of each group exhibit, by contrast, systematically distinct prosodic structures: the coordinated words in (406a) are formed by two prosodic words, whereas the coordinated words in (406b) are formed by a single prosodic word. The prosodic structure found in each group is exemplified in (407).

252 *Compound prosodic words*

- (407) a. (alegre)_ω (~~mente~~)_ω e (triste)_ω (mente)_ω >
 (alegre)_ω e (triste)_ω (mente)_ω
 b. (acampamento)_ω e (acantonamento)_ω >
 * (acampa)_ω e (acantonamento)_ω

Deletion under identity is also possible with morphological compounds, as illustrated in (408).

- (408) macro-economia e micro-economia > *macro e micro-economia*
 ‘macro-economy and micro-economy’
 foto-montagem e vídeo-montagem > *foto e vídeo-montagem*
 ‘photocomposition and videotape composition’
 luso-asiáticos ou afro-asiáticos > *luso ou afro-asiáticos*
 ‘Luso-Asiatic or Afro-Asiatic’
 auto-avaliação e hétero-avaliação > *auto e hétero-avaliação*
 ‘self-evaluation and evaluation made by others’
 inter-linguístico e intra-linguístico > *inter e intra-linguístico*
 ‘interlinguistic and intra-linguistic’

Again, DUI involving morphological compounds that are mapped onto a single prosodic word is impossible, as shown in (409) ³⁰⁵

- (409) automóvel e autocarro > * *automóvel e carro*
 ‘car and bus’
 telefone e telefonia > * *telefone e fonia*
 ‘telephone and radio’
 monografia e biografia > * *mono e biografia*
 ‘monograph and biography’

The prosodic contrast between compound words that allow/do not allow for DUI is illustrated in (410).

- (410) (auto)_ω (~~avaliação~~)_ω e (hétero)_ω (avaliação)_ω >
 (auto)_ω e (hétero)_ω (avaliação)_ω
 (monografia)_ω e (biografia)_ω >
 * (mono)_ω e (biografia)_ω

From this description, two facts should be retained: (i) no examples of DUI appear to exist where one of the complex words involved (or both) consist of a single prosodic word; (ii) the same morphological structures where DUI is attested, systematically yield bad results if the prosodic requirements on DUI are not met.

As said above, it is not our goal to determine all the conditions on DUI. Notice, in any event, that like in other languages (e.g. German – cf. Wiese 1996: 70), it appears that not all words that display the morphosyntactic and prosodic configurations just observed allow DUI to apply in EP.³⁰⁶ What is crucial for us is that those that allow DUI always seem to obey the phonological conditions of the construction: the complex words involved in the deletion process must show partial phonological identity, and both the element to be deleted and the element that remains in the string must be a prosodic word.³⁰⁷

The second process investigated in this section is clipping, a phenomenon that consists of the deletion of part of a word, while the meaning and the morphosyntactic properties of the original expression are maintained.³⁰⁸ As we have seen in chapter 1, the (minimal) prosodic word has been proposed to form a prosodic template for the output of the clipping process in several languages (cf. Prieto 1992, Cabré and Kenstowicz 1995, Thornton 1996 – see chapter 1, section 6.1). The following discussion reveals the importance of the prosodic word for clipping also in EP. Nevertheless, it appears that this constituent plays a role already in the input of the process.

As the examples in (411) show, clipping may affect various kinds of morphosyntactic expressions, such as those containing two words (see 411a), a stem and a word (see 411b), or a prefix-like element and a word (see 411c). Notice that, in the latter cases, the deleted element may either correspond to a word, a stem or a prefix.

- (411) a. *computador portátil* > *portátil* ‘portable (computer)’
e-mail > *mail* ‘id.’
JSD (jota esse dê) > *J* (abbreviation for a political party)
BMW (bê eme dâblío) > *BM* ‘id.’
cabelos brancos > *brancos* ‘white (hair)’³⁰⁹
paragem cardíaca > *paragem* ‘(heart) failure’³¹⁰
Ficheiros Secretos > *Ficheiros* ‘lit. (secret) files (*The X-Files*)’
Jogos Olímpicos > *Jogos* ‘(Olympic) Games’³¹¹

254 *Compound prosodic words*

b. <i>telemóvel</i>	> <i>móvel</i>	‘portable (phone)’
<i>hipermercado</i>	> <i>híper</i>	‘large (store)’
<i>micro-ondas</i>	> <i>micro</i>	‘micro(waves)’
<i>químio-terapia</i>	> <i>químio</i>	‘chemo(therapy)’ ³¹²
c. <i>internet</i>	> <i>net</i>	‘id.’
<i>homossexual</i>	> <i>homo</i>	‘homo(sexual)’
<i>heterossexual</i>	> <i>hétero</i>	‘hetero(sexual)’
<i>bissexual</i>	> <i>bi</i>	‘bi(sexual)’

Despite the diversity in the morphosyntactic composition of the expressions subject to clipping, it invariably involves the deletion of one (or more) of two (or more) prosodic words that compose the original expression, as shown in (412).³¹³

(412) (computadOr) _ω (portÁtil) _ω	>	(portátil) _ω
(mIcro) _ω (ONdas) _ω	>	(micro) _ω
(INter) _ω (nEt) _ω	>	(net) _ω

Thus, in EP, unlike in the other languages mentioned above (see also the Italian examples in chapter 1, section 6.1), clipping only applies to words composed of more than a single prosodic word, and the clipped form must correspond to one of the prosodic words of the input expression.³¹⁴

In short, like coordination reduction, clipping also seems to depend on the presence of two prosodic words in the expression that forms the input of the process. These phenomena, therefore, provide additional evidence for the prosodic word domain in EP. We should remark that neither deletion under identity nor clipping seem dependent on the existence of a compound prosodic domain. Although deletion under identity appears to involve only constructions that correspond to compound prosodic words, there are certain types of compound prosodic words that do not undergo this kind of reduction (see note 307). As for clipping, it may apply to combinations of words that apparently do not form compound prosodic words (as most of the examples in 411a above).

To sum up, the two phenomena considered in this section support the analysis developed in this chapter and in the preceding one as far as the (minimal) prosodic word domain is concerned. Deletion under identity and clipping refer to constituents that are not defined on morphosyntactic but rather on prosodic terms. The relevant constituents have the properties of

prosodic words, as defined throughout this book on the basis of other phonological criteria.

5. Discussion

In this chapter we have seen that a number of constructions composed of more than one prosodic word display a phonological behavior distinct from other combinations of prosodic words within phonological phrases. In order to account for their phonological specificities, we have proposed that the elements belonging to these constructions are grouped together within a compound prosodic word domain. This final section is devoted to the identification of the common features shared by all these constructions which may explain their specific prosodization.

Given that the compound prosodic word is not a new prosodic constituent, but rather a constituent of the same type as the prosodic word, we put forward the hypothesis that the same general mechanism that is responsible for the formation of the prosodic word domain is also responsible for the formation of compound prosodic words. Nevertheless, as we will see, some additional information must be included in the proposal outlined in chapter 5, so as to account for the various kinds of constructions studied in the present chapter. This information has to do exclusively with the relation between word stress and the prosodic word. In fact, all the constructions surveyed in chapter 5 have a single word stress, which has been assumed in the EP literature to be assigned to a unit containing a stem (see chapter 3, section 2). Given that the prosodic word was also proposed to be built with reference to the stem, there is a systematic relationship between a word stress and a prosodic word. In this chapter, however, we have seen that in some cases there are stressed units that do not include a stem. Thus, something more has to be said concerning the relation between word stress and the prosodic word domain. Before doing so, a discussion on the sources of word stress in EP is required.

5.1. *On the sources of word stress*

In many languages, including EP, a stem plus (some) following suffixes always form a prosodic word, that is a sequence which bears one primary word stress (e.g. Italian, Greek, Latin, Hungarian, Dutch, Sanskrit, French, Spanish – cf., among others, Nespor and Vogel 1986 for several of these

256 *Compound prosodic words*

languages, Hannahs 1995b for French, Peperkamp 1997a for Italian and Spanish). Other morphological units, by contrast, seem to be prosodized in different ways. In some languages, this depends on the segmental string. For example, German suffixes starting with a consonant are claimed to form an independent prosodic word, whereas suffixes starting with a vowel belong to the same prosodic word as their bases (cf. Wiese 1996). In Dutch, most of the suffixes that show the segmental and syllabic properties of prosodic words are suggested to form independent prosodic words, whereas suffixes that do not have such properties are included in the same prosodic word as their bases (cf. Booij 1995). In Italian and Spanish, prefixes are seen to form prosodic words if they are disyllabic (i.e. if they conform to minimality requirements – cf. Peperkamp 1997a). In all these cases, it appears that the presence of word stress may be inferred from the prosodic word status of the relevant morphemes.

Unlike in the languages mentioned above, in EP neither the segmental string nor the syllabic composition of the morphological units correlate with their prosodic status. This can be seen in (413), where a comparison between stressed and unstressed morphemes is presented (recall that [i] and [ɐ] correspond to underlyingly unreduced vowels).

(413) a. Stressed prefixes		Stressless prefixes	
<i>pré-</i>	[prɛ́]	<i>re-</i>	[Rɪ]
<i>pós-</i>	[pɔ́ʃ]	<i>des-</i>	[diʃ]
b. Suffix-like units with a stress independent of the base		Suffixes that form a stress domain with their base	
<i>-zinho</i>	[zɨ́nu]	<i>-dade</i>	[dádi]
<i>-mente</i>	[mẽ́ti]	<i>-mento</i>	[mẽ́tu]
<i>-avo</i>	[ávu]	<i>-eza</i>	[éʒɐ]
c. Stressed monosyllabic words		Stressless monosyllabic words	
<i>dê</i>	[dɛ́] ‘give’	<i>de</i>	[di] ‘of’
<i>nu</i>	[nú] ‘naked’	<i>no</i>	[nu] in-the-MASC
<i>já</i>	[ʒá] ‘already’	<i>da</i>	[dɐ] of-the-FEM
<i>i</i>	[í] (name of the letter i)	<i>e</i>	[i] ‘and’

d. Stressed disyllabic words			Stressless disyllabic words		
<i>parque</i>	[párki]	'park'	<i>porque</i>	[purki]	'why'
<i>pára</i>	[párɐ]	'(he) stops'	<i>para</i>	[pɐrɐ]	'to'
<i>onde</i>	[õdi]	'where'			

With respect to the contrasts in (413c) and (413d), the presence/absence of word stress correlates to some extent with the lexical/function word distinction: while content words (which include at least a stem) always bear word-stress, in the case of function words, some are stressed (e.g. *onde*) and some are not (e.g. *porque*), regardless of their segmental composition. Given the similar segmental composition of the words displaying the different patterns, the presence/absence of stress assigned to function words must be accounted for by some lexical property associated to the relevant words.³¹⁵ What remains unclear is whether stressed function words are the exceptional cases within the function word class or the unstressed function words are the marked cases, instead. Since some information concerning the location of stress must in any event be present in the lexical entry of some function words, as shown by the unpredictability of stress location in the prepositions in (414), we propose that stressed function words with one or two syllables must be lexically marked for stress.

(414) Stress pattern in EP disyllabic prepositions

Penultimate stress			Final stress		
<i>ante</i>	'before'	<i>entre</i>	'between'	<i>após</i>	'after'
<i>contra</i>	'against'	<i>sobre</i>	'on'	<i>até</i>	'until'
<i>desde</i>	'since'				

The same apparently unruléd behavior emerges from the contrasts in (413a) and (413b) above, where examples are given of segmentally similar prefixes and suffixes that show differences with respect to the presence/absence of word stress. With respect to prefixes, we believe that the similarity between prefixes and stems, pointed out in section 3.2, may account for the fact that some prefixes are treated like stems for the purpose of stress assignment in EP. Notice that many productive stressed prefixes may also occur as independent words, as those in (415).³¹⁶

(415) <i>pró</i>	<i>entre</i>	<i>extra</i>
'in favor of'	'between'	'additional piece'

Regarding suffixes, we believe that it is possible to predict which units form stress domains independent of their base. Despite the segmental similarity between the suffixes illustrated in (413b) above, there is a morphological distinction that sets the two classes of suffixes apart: the suffixes that are included in the stress domain of their base attach to stems and themes, whereas the suffixes that form independent stress domains attach to words (see Villalva 1994 on these morphological notions, and chapter 5, section 6 of this book for a review). If we assume that primary word stress is in general assigned to the last syllable of the derivational stem domain (in the spirit of several proposals on word stress assignment in EP – see chapter 3, section 2), then the presence of two word stresses may follow from the morphological structure displayed by the relevant suffixed words.³¹⁷ According to Villalva (1994), *z*-evaluative suffixes as well as *–mente* adjoin to the morphological word node (MW) yielding a constituent of the same type.³¹⁸ Under this approach, these words form two domains for word stress, as illustrated in (416) (where a simplified version of the morphological structure assigned to similar examples in Villalva 1994: 12 is given).³¹⁹

- In the case of morphological and syntactic compounds, the stress pattern follows straightforwardly from their morphological composition: each derivational stem forms a domain for word stress assignment.

The final case we have to address is the stress pattern of mesoclititic structures (see chapter 4, section 4.3 and section 3.6 above). That the infinitive form of the verb bears word stress is not problematic, as the stress is obtained similarly to infinitive verbs in other syntactic configurations. As for the word stress assigned to *–haver*, we believe that it is one of the properties that allows this element not to be reinterpreted as a morphological affix. Our proposal is that this unit is phonologically treated like a(n auxiliary) verb in the lexical component, where it is assigned word stress. Notice that *–haver* displays precisely the same stress pattern as other verb forms whose stress is regularly assigned within the lexicon, as shown in (417) (see also chapter 2, section 4.1 and Pereira 1999).³²⁰

(417)	Future	Conditional	–haver (Future)	–haver (Conditional)
1Sg	<i>falarEI</i>	<i>falarIa</i>	<i>–EI</i>	<i>–Ia</i>
2Sg	<i>falarÁs</i>	<i>falarIas</i>	<i>–Ás</i>	<i>–Ias</i>
3Sg	<i>falarÁ</i>	<i>falarIa</i>	<i>–Á</i>	<i>–Ia</i>
1Pl	<i>falarEmos</i>	<i>falarÍamos</i>	<i>–Emos</i>	<i>–Íamos</i>
3Pl	<i>falarÃO</i> 'to speak'	<i>falarIam</i>	<i>–ÃO</i>	<i>–Iam</i>

To sum up, we have seen that in EP word stress is regularly assigned to the last syllable of a derivational stem (cf. Mateus 1983, and the literature referred to in chapter 3, section 2). This assumption accounts for the stress pattern of (i) simple content words, (ii) derived words containing suffixes that attach to stems or themes, (iii) derived words containing suffixes that attach to words, (iv) morphological compounds, (v) syntactic compounds, and (vi) mesoclititic structures (since for the purpose of stress assignment, *–haver* appears to be still treated like a full verb). There are only two classes of cases left out, which behave asymmetrically: (i) within the closed class of function words, some elements are stressed and some are not; (ii) the same happens with prefixes. In the former case, we have assumed that in order for mono- or disyllabic function words to be stressed they must be lexically marked; in the latter case, we have proposed that some prefixes are treated like stems for the purpose of stress assignment, while other prefixes are not, in which case they fail to bear word stress.³²¹

5.2. *On the relation between word stress and the prosodic word*

Having determined the sources of word stress, we may now discuss the relation between word stress and the prosodic word domain. In all the cases under analysis, we have assumed that word stress in EP is assigned independent of the prosodic word domain. The prosodic word domain, in turn, is basically constructed with reference to the stem. Nevertheless, as we have seen, word stress may also be assigned to a unit that does not include a stem, which forms an independent prosodic word. This may be seen to be the case of stressed prefixes, as well as of the unit that contains the rightmost stress in suffixed words with *z*-evaluative suffixes, *–mente* and *avos*. In order to allow such stressed units to form prosodic words, we introduce the general condition stated in (418), concerning the relation between word stress and the prosodic word domain.

- (418) A unit bearing word stress must be included within
a minimal prosodic word

This condition seems to be operative both lexically and postlexically. We have evidence that stressed prefixes are subject to lexical rules that apply at the edge of the prosodic word domain (e.g. the lowering of vowels in syllables closed by /r/ – see chapter 3, section 8), and thus they must form prosodic words already in the lexicon. Given that we have proposed in chapter 5 that the right brackets of lexical prosodic words are not projected postlexically, we also need this generalization to apply postlexically in order to derive the prosodic word status of affixes that form prosodic words distinct from the prosodic words that dominate their morphological bases (we will come back to this point in section 5.3).

We must further clarify the prominence relations within compound prosodic words. As seen in chapter 1, section 6.1, the prosodic word is assumed by many researchers to include one and only one word stress. Consequently, so as to avoid the problem of having (compound) prosodic words with more than a single word stress, we propose a split in the well-formedness condition on the prosodic word related to the presence of word stress (already stated in 28 of chapter 1), as in (419) below.³²²

- (419) Well-formedness condition on the prosodic word domain
- a. A minimal prosodic word has one and only one
(word) primary stress

- b. A maximal prosodic word has one and only one prominent element

As it stands, the well-formedness condition on prosodic words allows the existence of compound prosodic words, since these are composed of minimal prosodic words that bear one and only one word stress that are grouped together under a maximal prosodic word, which has a single prominent element. It allows, furthermore, the most common situation where the notions of minimal and maximal prosodic word coincide. In such cases, there is a single word stress that is the prominent element of the (non-compound) prosodic word.

Finally, the prominence relations within the compound must also be defined. As we have seen, the head of compound prosodic words in EP always corresponds to the rightmost prosodic word. This is stated in (420).

(420) Compound prosodic word prominence (EP)

The head of a compound prosodic word is its rightmost prosodic word.

Besides the EP data pointing to the existence of a prominent element within the compound prosodic word, some facts from other languages also suggest that the prominence within (what we may assume to be) prosodic compounds must be stated independently of, for instance, phonological phrase prominence. This is shown for instance by the stress pattern displayed by morphosyntactic compounds in languages like Dutch and English (e.g. Booij 1995; Nespor 1999b; Visch 1999). Similar to other right recursive languages, the phonological phrase in English is expected to show a weak-strong pattern (cf. Nespor and Vogel 1986). Nevertheless, it is well-known that compound words in English may exhibit a strong-weak pattern, regardless of the lexical category of their members, as illustrated in (421) (taken from Nespor 1999b: 138).

(421) Strong-weak pattern in English compounds

NOMINAL COMPOUNDS	<i>bláckbird</i>	<i>rádio station</i>
ADJECTIVAL COMPOUNDS	<i>cólorblind</i>	<i>séasick</i>
VERBAL COMPOUNDS	<i>áir condition</i>	<i>týpe write</i>

In Dutch, Visch (1999: 3.5) observes that nominal phrases are stress-final, whereas nominal compounds are stress-initial. Thus, both in English and Dutch, a compound word may display a stress pattern different from the one found within the phonological phrase domain. This amounts to say that the prominence relations within a compound must be defined independently of ϕ -level prominence. Under the present proposal, this may be accomplished by assuming that the relevant compounds are mapped onto a compound prosodic word and that the prominent element of this constituent is the initial prosodic word.³²³

Interestingly, the Dutch strong-weak pattern found in some compounds also characterizes complex words that are not morphosyntactic compounds but that are composed of more than one prosodic word (e.g. derived words with suffixes that form independent prosodic words) (cf. Booij 1995: 5.4). This suggests that, like in EP, the internal prosodic words of these constructions are prosodized as compound prosodic words, instead of simply being grouped together within a phonological phrase.³²⁴ Both the EP facts and the Germanic languages data just mentioned may be seen to indicate that the members of (some) word compounds, as well as of derived words with two internal prosodic words, are grouped together under a compound prosodic word, rather than a phonological phrase.³²⁵

To conclude, we should point out that according to Hannahs (1995b) certain groupings of prosodic words in French also behave differently from words directly combined within the phonological phrase. Again, in our terms these prosodic words would be grouped within a compound prosodic word, given that we assume, contrary to Hannahs, that the clitic group is not a domain of the prosodic hierarchy. A similar analysis could also be extended to Baule, where some units that contain internal prosodic words are reported to undergo rules that refer to the prosodic word domain (see Leben and Ahoua 1997, and chapter 1, section 4).

5.3. Definition of the prosodic word domain in EP: final remarks

We close this chapter with a summary of our main findings (presented in this chapter and in chapter 5) on the formation of the prosodic word domain in EP.

Like other prosodic domains above the word level (namely, ϕ and I – see, among others, Nespor and Vogel 1986, and for EP, Frota 2000), the prosodic word is obtained via a set of mapping relations, and is also subject to pure phonological conditions. The former are given in (422) (see also

chapter 5, section 6). In EP, the latter appear to be related only to stress facts, as shown in (423) (see also section 5.2), where the definition of *maximal prosodic word* and *minimal prosodic word* is the one repeated under (424) for ease of reference (see section 3.1).

(422) Mapping relations

Morphology–Phonology

The prosodic word domain includes a stem plus suffixes
(and non-transparent prefixes)

Lexical prosodic word–Postlexical prosodic word

(Only) The left edge of a lexical prosodic word is projected
postlexically

Syntax–Phonology

The lowest Lex^o is mapped onto a prosodic word

(423) Well-formedness conditions on the prosodic word domain

- A minimal prosodic word has one and only one (word) primary stress
- A maximal prosodic word has one and only one prominent element
- A unit bearing word stress must be included within a minimal prosodic word

(424) *Maximal prosodic word*: a prosodic word that is immediately dominated by the next higher prosodic level (i.e. the phonological phrase);

Minimal prosodic word: a prosodic word that immediately dominates the next lower prosodic level (i.e. the foot).

A specific clause concerning the elements that are left undominated by the prosodic word node was also seen in chapter 5, section 6 to be required in order to account for the prosodization of EP stressless prefixes and clitics. This is stated in (425).

(425) Prosodic parsing at the prosodic word level

Any morpheme not dominated by the prosodic word node is adjoined to the following prosodic word (if any).

Notice that the relation between word stress and the prosodic word as formulated in (423) is potentially universal. Additionally, the morphology-phonology and syntax-phonology mapping relations defined in (422) are very similar to those reported for other languages. For example, our morphology-phonology mapping condition is similar to the clauses of prosodic word formation proposed in Nespor and Vogel (1986) for several languages or in Hannahs (1995b) for French. Additionally, compound prosodic words in languages like Baule may be obtained via a syntax-phonology mapping condition like the one stated in (422) – although, according to the data in Leben and Ahoua (1997), it is X' rather than X^0 that is mapped into a prosodic word in this language.

By contrast, the relation between lexical and postlexical prosodic words seems to vary crosslinguistically. For example, while the right boundary of lexical prosodic words is not postlexically projected in EP (as shown by the sequences of verb plus enclitics), it appears to be so in languages like Neapolitan, since enclitics are adjoined to the preceding prosodic word (rather than incorporated into it) (see the data in Peperkamp 1997a, also presented in chapter 1, section 6.3 and chapter 5, section 4.3.1 of this book).

Besides this, the way languages prosodize elements that are not dominated by a prosodic word node also appears to vary greatly. This is particularly evident in the prosodization of clitics. As we have seen in chapter 1, section 6.3, languages vary in the way clitics are prosodically related to their hosts (i.e. clitics may be incorporated into or adjoined to the prosodic constituent that dominates their host), in the definition of clitic hosts (the prosodic word, ϕ or I), as well as in the direction of cliticization (i.e. a language may show a tendency for rightward or leftward cliticization and allow/do not allow bidirectional clitics). Therefore, (425) should be specific of languages like EP.

Finally, the phonological conditions on the prosodic word other than those related to word stress may also be language particular. For example, the definition of a minimal size for the prosodic word is active in many languages mentioned in chapter 1, section 6.1, but not in EP, as shown in chapter 5, section 2.

Evidence was also found for the existence of prominence relations within the compound prosodic word (see sections 3 and 5.2). Given the definition of compound prosodic words in (426) (stated above under 354), the prominence relations within phonological compounds in EP were stated as in (427) (already presented under 420). In this respect, compound prosodic words pattern like other postlexically built prosodic domains.

(426) *Compound prosodic word*: a prosodic word that dominates two (or more) constituents of the same type (i.e. prosodic words).

(427) Compound prosodic word prominence

The head of a compound prosodic word is its rightmost prosodic word.

Again, the prominence relations within compound prosodic words, similar to prominence patterns of other prosodic domains, may vary across languages. As we have seen above, in some Germanic languages the prominent element of certain morphosyntactic compounds and derived words is the initial prosodic word. This can be seen to be a consequence of the prosodization of these constructions as compound prosodic words and of the assignment of compound prominence to the leftmost prosodic word in these languages, unlike in EP.

We will now briefly demonstrate how the generalizations on the prosodic word domain summarized above are capable of predicting the prosodization of the constructions investigated in this chapter. We recall that the constructions claimed to be formed by two or more prosodic words that are grouped together within a compound are (i) derived words with suffixes that form stress domains independent of their base, (ii) derived words with stressed prefixes, (iii) morphological (stem plus stem) compounds, (iv) morphosyntactic and some syntactic (word plus word) compounds, (v) abbreviations, (vi) mesoclititic structures, and (vii) sequences of prosodic words consisting of (a) pairs of letter names, (b) names of letters followed by numerals, as well as (c) some numerals followed by the very frequent words *horas* and *anos*.

Let us start by considering the prosodization of derived words with suffixes that form stress domains independent of their base. An illustrative example is provided in (428) for a word like *claramente* ‘clearly’. Given the morphological structure of the derived word, *clara* is mapped onto a prosodic word (see 428i). This prosodic word, however, does not include the suffix *–mente*, since the suffix is a stressed unit and each minimal word may only contain one word stress. Thus, only the unit *clara*, and not *claramente*, forms an independent prosodic word as a consequence of the morphology-phonology mapping. As a stressed unit must correspond to a prosodic word, *–mente* forms its own prosodic word (see 428ii). Postlexically, however, only the left-edge of lexical prosodic words is projected (see

266 *Compound prosodic words*

428iii).³²⁶ Nevertheless, the presence of two stress domains, implies the formation of two prosodic words (see 428iv).³²⁷ Because these two prosodic words form a unit under a single Lex^o, and given that a Lex^o is mapped onto a prosodic word, a compound prosodic word results (see 428v).

(428) Derived words with suffixes that form independent stress domains

claramente 'clearly'

Lexical prosodization

i. (clara)_ω [clar: stem; -a: theme index]

ii. (mente)_ω

Postlexical prosodization

iii. ω(clara ω(mente

iv. (clara)_ω (mente)_ω

v. ((clara)_ω (mente)_ω)_ω^{max}

Derived words with stressed prefixes display a similar picture, as shown by the example in (429). A stem and the following suffixes (forming a single stress domain) are prosodized as a prosodic word (see 429i). The unparsed prefix acquires the prosodic word status by virtue of bearing word-level stress (see 429ii). Postlexically, only the left-edges of lexical prosodic words are projected (see 429iii). The stress on the units to the right of these boundaries entails their postlexical prosodic word status (see 429iv). As the derived word corresponds to a Lex^o-level unit, the whole structure is mapped onto a prosodic word (see 429v). However, given that this expression already contains two prosodic words, the new prosodic word has the properties of a compound prosodic word.

(429) Derived words with stressed prefixes

pós-sintáctico 'post-syntactic'

Lexical prosodization

i. (sintáctico)_ω [sintact: stem; -ic: suffix; -o: theme index; 0: inflection]

ii. (pós)_ω

Postlexical prosodization

- iii. ${}_{\omega}(\text{pós})_{\omega}(\text{sintáctico})$
- iv. $(\text{pós})_{\omega}(\text{sintáctico})_{\omega}$
- v. $((\text{pós})_{\omega}(\text{sintáctico})_{\omega})_{\omega}^{\text{max}}$

Before proceeding, it should be noticed that, although the steps in (429iii) and (429iv) may appear redundant, they are in fact relevant. Indeed, (429iii) could not be subsumed by (429iv) because the information concerning the morphological structure of the word is no longer available postlexically. Thus, despite the presence of two word stresses (that indicate the presence of two prosodic words), if it is simply assumed that a stressed unit must be contained within a minimal prosodic word, it is not possible to know where the first one ends and the second one begins (see the possibilities in 430 for a simple example where only one unstressed syllable occurs between the two stressed syllables). Conversely, it is not possible to do without step (429iv) because in that case the prefix could simply be adjoined to the prosodic word that includes its base, and thus the fact that it behaves like a prosodic word postlexically would not be accounted for.

- (430) $*(\text{pós sin})_{\omega}(\text{táctico})_{\omega}$
 $(\text{pós})_{\omega}(\text{sintáctico})_{\omega}$

Morphological compounds are the next case we will consider. According to Villalva (1994), morphological composition always consists of the concatenation of two stems. As mentioned in section 3.3, this author provides arguments against the distinction between stem+stem compounds and stem+word compounds in EP. Consequently, the difference between morphological compounds that form a single prosodic word and morphological compounds that form two (or more) prosodic words appears not to be due to a different morphological structure in EP (unlike proposals for Greek and Italian in Nespor and Ralli 1996 and Nespor 1999b). Assuming the analysis of Villalva for EP, we propose that the prosodic distinction may result, instead, from the fact that some compounds have a morphological internal structure that is phonologically transparent, while others have a morphological structure that is phonologically opaque. Only in the former case are the internal members of morphological compounds mapped onto independent prosodic words. That the morphosyntactic structure of complex words may become opaque is well attested both in EP and in other languages: for example, in prefixed words in Italian (cf. Nespor and Vogel

1986, Peperkamp 1997a), in Dutch (cf. Booij 1995: 120), and in French (cf. Hannahs 1995a: 2.6, for words derived with *in-* when attached to adjectival bases), as well as in compounds in Italian (e.g. Peperkamp 1997a, for familiar compounds). In EP, we have seen that some prefixed words may be treated like simple words (see chapter 5, section 3.2 and section 3.2 above); and the same may happen with morphological compounds (see section 3.3), syntactic compounds (see section 3.4), and abbreviations (see section 3.5). The fact that newly created morphological compounds typically form two stress domains in EP is in accord with the transparent/opaque distinction.

Given the generalizations and principles stated above, this proposal allows us to straightforwardly account for the prosodization of morphological compounds. Morphological compounds whose internal structure is phonologically opaque are prosodized similar to words formed with a single stem, that is, as single prosodic words. As for morphologically transparent compounds, they surface with more than a single prosodic word, since each stem forms a prosodic word with any following suffix (when itself does not form an independent prosodic word).

The illustration of the prosodization procedure of transparent morphological compounds is given in (431). The morphology-phonology mapping relation yields the prosodic word status of both *macro* and *económico* (see 431i). Postlexically, only the left boundaries of lexical prosodic words are projected (see 431ii). The units to the right of the postlexically projected brackets, however, bear word stress and thus form independent prosodic words (see 431iii). Being a Lex^o-level expression, the morphological compound is prosodized as a single prosodic word (431iv). As this prosodic word already contains two minimal prosodic words, a compound prosodic word is formed.

(431) Morphological compounds

macro-económico 'macro-economic'

Lexical prosodization

- i. (macro)_ω (económico)_ω [*macro*: stem; *econom*: stem;
–*ic*: suffix; –*o*: theme index;
0: inflection]

Postlexical prosodization

- ii. _ω(macro _ω(económico
iii. (macro)_ω (económico)_ω
iv. ((macro)_ω (económico)_ω)_ω^{max}

The prosodization of word plus word compounds is also explained by the same general mechanisms. Here too, the relevant piece of information that may account for the phonological difference among word plus word compounds is the morphosyntactic transparency/opacity of the internal syntactic structure of the compound. If we assume, with Villalva (1994) (see section 3.4 above), that only V+N compounds and other lexicalized compounds do not have a transparent syntactic structure, then only the compounds of this type form a single Lex⁰-level expression. Other word plus word compounds, by contrast, have a transparent syntactic structure, and therefore each member of these expressions is a Lex⁰-level expression.³²⁸ This distinction has immediate consequences for the prosodic grouping of the prosodic words that form the different compounds. If the whole compound forms a single Lex⁰-level expression, the whole expression is expected to be mapped onto a prosodic word. Given that this prosodic word contains internal prosodic words, it must form a compound prosodic domain. Instead, if each member of the compound is a Lex⁰-level expression – that is, if the compound has a complex syntactic structure – then the prosodic words that integrate the syntactic compound are not grouped within a compound prosodic word. We illustrate each of these cases under (432) and (433), respectively (except for the absence of step (iv) in 433, the procedure is similar to the one just described for morphological compounds).

(432) Morphosyntactic (V+N) compounds

abre-latas ‘can opener’

Lexical prosodization

- i. (abre)_ω (latas)_ω [*abr*: stem; *-e*: theme vowel; 0: inflection; *lat*: stem; *-a*: theme index; *-s*: inflection]

Postlexical prosodization

- ii. _ω(abre _ωlatas
 iii. (abre)_ω (lata)_ω
 iv. ((abre)_ω (latas)_ω)_ω^{max}

(433) Syntactic compounds

surdo-mudo ‘deaf and dumb’

270 *Compound prosodic words*

Lexical prosodization

- i. (surdo)_ω (mudo)_ω [*surd*: stem; –*o*: theme index; 0: inflection;
mud: stem; –*o*: theme index; 0: inflection]

Postlexical prosodization

- ii. _ω(surdo) _ω(mudo)
 iii. (surdo)_ω (mudo)_ω

Let us now consider the prosodization of mesoclitic structures. Recall that, according to the analysis of Duarte and Matos (2000), mesoclitic constructions are formed by the infinitive form of the verb, a clitic pronoun and a “lexicalized T-affix” (see chapter 4, section 4.3). Regardless of the exact nature of the later constituent, Duarte and Matos’ analysis suggests that it does not qualify as a syntactic Lex^o, in contrast with the infinitive verb. Nevertheless, according to the same proposal, the affix is attached to the functional head that constitutes the syntactic landing site of the verb plus clitic sequence (which is Tense). Therefore, it is the whole functional head that includes the verb that functions as a Lex^o. Furthermore, the affix-like element bears its own stress, which we assume to have been assigned at the lexical level, in a way similar to other (auxiliary) verbs (see section 5.1). Given this analysis, the prosodization of mesoclitic expressions proceeds regularly, as illustrated in (434). The verb stem and the following affixes form a single prosodic word at the lexical level (see 434i). Because it lacks word stress, the pronominal clitic (which does not include a stem) is lexically prosodized as a syllable but not as a prosodic word (see 434ii). Notice that, unlike prefixes, it may not adjoin to the next prosodic word at this point because clitics are not combined with their hosts at the lexical level. The inflected form of the auxiliary-like element *–haver* is prosodized as a prosodic word, since it is stressed (see 434iii). Postlexically, only the left boundaries of the lexical prosodic words are projected (see 434iv). Given the existence of two stress domains, two prosodic words are then built (see 434v). The fact that the whole mesoclitic expression is dominated by a single Lex^o causes its prosodization as a single prosodic word (see 434vi). This prosodic word contains two minimal prosodic words, and thus it forms a compound prosodic word.

(434) Mesoclitic structures

pedir-lhe-ia ‘(I) would ask him’

Lexical prosodization

- i. (pedir)_ω [ped: stem; -i: theme vowel; -r: inflection]
- ii. lhe
- iii. (ia)_ω

Postlexical prosodization

- iv. ω(pedir lhe ω(ia
- v. (pedir lhe)_ω (ia)_ω
- vi. ((pedir lhe)_ω (ia)_ω)_ω^{max}

The last structures to be examined are those formed by letter names and/or numerals. As for the internal members of these structures, their prosodic status follows from their lexical properties. Under the assumption that they are composed of a stem and a theme index, each element of the expression forms its own prosodic word. As for the prosodization of the whole expression within a (compound) prosodic word, our hypothesis is that these expressions do not have an internal syntactic structure, so that the whole abbreviation or other similar expressions constitute a single Lex⁰-level constituent.

The prosodization of abbreviations and other combinations of letter names, which are assumed to be inserted under a single Lex⁰, is illustrated in (435).

(435) Abbreviations

LM (ele eme) ‘a tobacco brand’

Lexical prosodization

- i. (ele)_ω (eme)_ω

Postlexical prosodization

- ii. ω(ele ω(eme
- iii. (ele)_ω (eme)_ω
- iv. ((ele)_ω (eme)_ω)_ω^{max}

We should point out that, while abbreviations are always mapped onto a single prosodic word (that contains two or more minimal prosodic words), other expressions show variation within the same morphosyntactic category: for example, the combination of a numeral like *onze* with a noun like *horas* yields a compound prosodic word, and thus under the present pro-

posal this combination should count as a Lex⁰-level expression; however, the combination of a numeral like *sete* with the same noun *horas* does not behave like a compound, and thus each element should count as a Lex⁰-level expression. These specificities must follow from some lexical property of particular sequences of items. Our hypothesis is that these expressions have lexicalized as a single Lex⁰, and this is why they are mapped onto a compound prosodic word (see section 3.7).

To conclude, with a small set of generalizations and conditions, we were able to account for the prosodization of prefixed and suffixed words forming two stress domains, morphological, morphosyntactic and syntactic compounds, as well as mesoclititic structures and abbreviations. A subset of the same generalizations and conditions was also shown in chapter 5 to account for the prosodization of prefixed and suffixed words that form a single stress domain, as well as for the prosodization of enclitic and proclitic function words in EP.

²⁷³ Before proceeding, we should point out that it is not always possible to determine the behavior of a given construction with respect to all relevant phenomena. In fact, in some cases, no examples can be found where the context for a given rule either to apply or to be blocked occurs.

²⁷⁴ On the morphological status of *z*-evaluative suffixes and *-mente*, see, for example, Villalva (1992, 1994). Despite the fact that these units show morphosyntactic properties of both suffixes and words, which makes their classification difficult to establish, we will refer to these elements as suffix(-like) units, following the traditional terminology. In the translation of *z*-evaluative suffixes, we use DIM for diminutive and AUG for augmentative.

²⁷⁵ We use the expression *word stress* to indicate that each element of these constructions bears a word-level stress, instead of a secondary rhythmic stress. This does not preclude, nevertheless, that one of the stressed elements is the prominent element of the whole construction, as we will see further below. For the demonstration that the stress on the base is also assigned with reference to morphological information see, for example, Pereira (1999), who calls this stress “morphological secondary stress”.

²⁷⁶ In Cunha and Cintra (1984) *avos* is written as a separate word. We will adopt this representation, despite the fact that we will assume this element to have the same status as *z*-evaluative suffixes and *-mente*.

²⁷⁷ Inflection is only visible with *z*-evaluative words, since the morphological base in this case may inflect for plural. As for *-mente*, it attaches to adjectival bases that do not inflect for plural, and, although it selects a feminine adjectival base, according to Villalva (1994: 4.2.2.1) gender cannot be considered to be inflectionally marked in EP. Finally, *avos* attaches to numerals that are invariable. Recall, to conclude, that inflection may be marked by a phonetically empty constituent (cf. Villalva 1994, and section 5.5).

²⁷⁸ To be exact, the central vowel is not lowered, since it is already low underlyingly. Nevertheless, this vowel patterns like the other vowels in prosodic word final syllables closed by sonorants in that it does not reduce.

²⁷⁹ The example in (11) is adapted from Frota (2000a) (the prosodic bracketing is ours). The syllable bearing the focal stress is signaled with capital letters in boldface. Recall that the assignment of focal stress affects I-level prominence and implies the compression of the range immediately after the focal stress (see section 3.18 for further details). This is an important remark since the first syllable of the word may also bear emphatic stress, which has different phonological and semantic/pragmatic consequences (see also section 3.16).

²⁸⁰ Recall that focal stress is always assigned to the most prominent element of the relevant *word*, which in non-compound words is the syllable bearing word stress.

²⁸¹Andrade (1984: 23) reports that the stressed syllable of the base in words with *z*-evaluative suffixes shows a duration significantly lower than the stressed syllable of the suffix. However, as pointed out by the author, the constructions studied occur in phrase final position. The different values may thus follow from the presence of phrasal prominence on the last prosodic word of the constituent, rather than from the prominence relations within these words.

²⁸² Notice that this does not preclude that a pitch accent may also appear associated to the first prosodic word of the compound, when the second prosodic word also has a pitch accent.

²⁸³ Two facts point to the relevance of the notion *head of the compound* for the blocking of non-back vowel deletion: (i) the first one is the similarity in the blocking contexts of all processes of vowel deletion – regardless of the specific prosodic domain, vowel deletion always seems to be obligatorily blocked when the target vowel is followed by a vowel belonging to the prominent prosodic word of the constituent it is part of; (ii) the second fact concerns the behavior of vowel deletion in compound words (like those formed by some abbreviations – see section 6.2.5) that include more than two prosodic words. In the latter cases, vowel deletion is not obligatorily blocked when the second vowel is stressed and belongs to an internal prosodic word, but it is obligatorily blocked when V2 is stressed and belongs to the head of the compound.

²⁸⁴ This term should not be confused with the term *minimal word*, used to refer to the minimum size of prosodic words in several languages (see section 1.5.1).

²⁸⁵ We believe that there is some inter- and intra-speaker variation concerning the possibility of round vowel deletion when two prosodic words belong to the same ϕ but the head of the constituent is not involved. This variation appears to be dependent on register, since vowel deletion is felt to be more likely to occur in more informal registers. What is crucial is that, even in these informal styles, the deletion of the vowel within compounds is totally excluded, and thus a clear distinction emerges between a sequence of two prosodic words within the ϕ -phrase and within the compound prosodic word.

²⁸⁶ We have not collected data on the possibility of assigning initial stress to these structures, but it seems clear to us that it can also occur in this context, like emphatic stress.

²⁸⁷ In the prosodic tree, *w* and *s* stand for ‘weak’ and ‘strong’, respectively. As already pointed out, the relevance of prominence relations within the prosodic compound will become clear later on in this chapter.

²⁸⁸ Despite the similarities between prefixes and some stems in morphological compounds, there is at least one test that appears to set these categories apart in EP: non-high non-central vowels in final position of an internal stem of transparent morphological compounds always surface as low (cf. *tél[ɛ]-chamada* ‘lit. distant

call'; *mon[ɔ]-acentual* 'monoaccentual'). Given that this behavior is specific of compound internal stems, examples of reduced final unstressed vowels are expected to be found with prefixes. Although rare, examples of this kind do exist (cf. *vice-*, where the final non-back vowel undergoes vowel reduction and is usually deleted, like similar underlying vowels in prosodic word final position). Interestingly, this morpheme may show a low non-back vowel for some speakers, thus suggesting that these speakers treat it like a stem rather than a prefix.

²⁸⁹ There is dialectal variation with respect to the realization of this type of words. Nevertheless, it is well-known that in the Lisbon variety we are describing [ɐjʃ] is not a natural realization in such non-transparent prefixed words, in contrast with the realization of the prefix in transparent constructions.

²⁹⁰ This apparently differentiates EP from Italian (see section 1.5.5). According to Nespor and Vogel (1986), monosyllabic prefixes in Italian, in transparent constructions, form prosodic words independent from their bases if they end with a vowel, but they are incorporated into the prosodic word that dominates their base if they end with a consonant. Adopting a different approach, Peperkamp (1997a) claims that in Italian and Spanish monosyllabic prefixes never form independent prosodic words, since they do not observe the minimality requirement assumed to be imposed on prosodic words. In French, by contrast, monosyllabic prefixes are argued to form prosodic words (cf. Hannahs 1995b).

²⁹¹ It is not possible to test the behavior of these constructions with respect to round vowel deletion and non-back vowel deletion because we could only identify one prefix ending with the relevant vowels (namely, *vice-*) and there are no words where it is followed by a stressed vowel. With respect to syllable degemination, we only found one case where the context for the rule could be met (namely, *vice-secretário* 'vice-secretary'). We do not think syllable degemination is likely to apply here because the consonants involved are fricatives (but see section 3.14, for other descriptions according to which syllable degemination is possible in this context). In any event, we judge similarly degemination in this segmental context in non-compound environments (e.g. *disse segredos* '(he) told secrets'). Thus, these facts cannot be taken to mean that syllable degemination does not apply within prefixed words formed by two prosodic words. Notice that syllable degemination has been suggested not to occur with prefixes in general (cf. Villalva 1994: 157). Nevertheless, the examples given by Villalva involve either stressless prefixes (which are expected not to undergo the process because they do not form independent prosodic words), or a segmental string that does not display the context for the rule to apply (as in the example *super-perto* 'very closed' – see section 3.14 for the description of this process).

²⁹² As for initial stress and emphatic stress, we have no data concerning the possibilities of assigning these types of stresses to the initial position of the prosodic word that includes the morphological base of stressed prefixes. The same

is true for the remaining constructions studied in this chapter. Since our intuitions in these cases are not sound enough, we cannot say whether initial stress and emphatic stress may be assigned to the internal members of compound prosodic words. This is an issue for further research.

²⁹³ It is often assumed that the compound linking vowels in the language are –i– and –o– (e.g. Villalva 1994: 6.1.2), the former appearing frequently at the end of the first member of compounds with Latin stems, and the latter occurring often at the end of the first member of compounds with Greek stems (e.g. Cunha and Cintra 1984). From a phonological viewpoint, the linking vowel –o– behaves like non-high non-back vowels and round vowels at the end of internal members of compounds, since all these vowels are realized obligatorily as low (e.g. *israEl*[ɔ]-*palestinIano* ‘israelian-palestinian’, *tÉl*[ɛ]-*chamAda* ‘distant call’, *BI*[ɔ]-*mÉdico* ‘biomedic’, respectively).

²⁹⁴ We should point out that many of the exceptional cases of absence of vowel reduction in stressless positions involve morphological compounds that have lexicalized as a single prosodic word. Examples of exceptionally unreduced stressless vowels are found in words such as *homorgânico* ‘homorganic’ and *homônimo* ‘homonym’, where the first stem is not perceived as stressed and it may count for the computation of the primary word stress. Notice that the initial vowels of these words are obligatorily realized as low, as usually happens with vowels that exceptionally fail to undergo vowel reduction.

²⁹⁵ As the first stems end in [i] or in a low vowel, we may not test the behavior of these compounds with respect to final vowel deletion processes or syllable degemination.

²⁹⁶ With respect to the form *grande área* there seems to be interspeaker variation as to its status as a compound: for some speakers its meaning is compositional, while for other speakers it is merely the name of a part of a football field. For the latter group of speakers, including us, the presence of the glide is obligatory. For the former group of speakers vowel deletion appears to be possible in this context.

²⁹⁷ The difficulty in judging these cases is probably related to the morphosyntactic analysis of this type of expressions. According to Villalva (1994: 6.2.2), speakers hesitate in considering A+N and N+A expressions as compounds given that their interpretation is similar whether they form X^0 or X^{\max} -level expressions, unless they have undergone a process of semantic lexicalization. In the examples under observation, the meaning of the expressions is very close to their compositional meaning. Our hypothesis concerning these structures is that the ambiguity of the morphosyntactic analysis may still have an effect on the prosodization of this type of compounds (which may either form a prosodic compound or not). This issue is left for future investigation.

²⁹⁸ In a reading task with 4 speakers, described in detail in chapter 7, this word was realized with a glide by two speakers (in two realizations – one at normal speech

rate and the other at fast speech rate), and with no glide in the two realizations of two other speakers. Although we find the glide possible when it occurs in phrase final position, as it was the case in the text read by the four speakers, we tend to prefer the realization where the vowel is deleted, and more so if the compound word is not in phrase final position, as in (44c). Again, the reasons for this type of variation require further investigation. We should point out, nevertheless, that, while left headed N+N compounds like *cobra-cascavel/cobras-cascavel* ‘rattlesnake(PL)’ allow for inflection in the first noun (cf. Villalva 1994: 381-382), *verde-água* – a left-headed A+N compound – does not seem to allow for inflection at all. Other words denoting colors that have a specific behavior with respect to inflection are *cor de rosa* ‘pink’ and *cor de laranja* ‘orange’. For Villalva (1994: 388), these formations must be considered lexicalized compounds or in a process of lexicalization.

²⁹⁹ According to their behavior with respect to a number of tests considered in Villalva (1994: 6.2.2) (including the (im)possibility of internal inflection, as well as morphological and syntactic modification tests), left headed N+N compounds, and A+A and N+N coordinated compounds are classified as syntactically opaque and morphologically transparent; and expressions composed of N+PP, A+N and N+A are claimed to be both syntactically and morphologically transparent.

³⁰⁰ There is only one letter of the alphabet that starts with a stressless vowel (*h* [ɐgá]), but we found no abbreviations with this letter. In the expression referring to the blood type *Rh+* or *Rh-* ([éɾɐgá]) there is (nearly) obligatory vowel deletion. This is the expected behavior, given that V2 is unstressed.

³⁰¹ An issue not treated in this section concerns the maximum number of prosodic words with a compound domain. Frota (2000a) suggests that compound intonational phrases are possibly restricted to a binary setting in EP. The data from abbreviations seems to indicate that compound prosodic words may be formed at least by three elements of the same type. In the case of abbreviations formed by four prosodic words, like *PCTP*, *MRPP* and *CNRT* (the former two are abbreviations of Portuguese political parties, and the latter is an abbreviation of an East Timor political party), it is possible to group the first two prosodic words together and then the last two prosodic words. This is shown by the realization of *MRPP* (*eme erre pê pê*, [émjérpépé]), where, for many speakers, including us, the non-back vowel in the first prosodic word may not be deleted. In this case vowel deletion is obligatorily blocked because the second prosodic word is the head of a prosodic word compound. By contrast, in *CNRT* (*cê ene erre tê*, [séénérté]), the blocking no longer applies. Under our view, this is so because the third prosodic word is not the head of the prosodic word compound that it forms with the following prosodic word. The data points to a binary setting of compound prosodic words in EP, although a ternary setting should also be allowed in structures with an odd number of prosodic words (as, for example, *RFM*).

³⁰² As mentioned in section 2.3.2, Van der Leeuw (1997) also assumes that pronominal clitics in mesoclitic structures are enclitic to the preceding prosodic word, although no empirical evidence for this analysis is provided.

³⁰³ Because the verb in the future and in the conditional forms does not end in a non-back vowel or in a round vowel, we may not assess the behavior of the vowel deletion processes involving these vowels in compound word final position.

³⁰⁴ Notice that an *h* at the beginning of written words in EP is purely conventional, as it never corresponds to an actual speech sound or an underlying segment.

³⁰⁵ The non-independent prosodic word status of the first stem is shown by the fact that its stressed vowel, as well as the linking vowel/the final vowel of the first stem undergoes vowel reduction (to the exception of *autocarro*, where the final vowel of the first stem exceptionally surfaces as low).

³⁰⁶ For example, in EP, it is not clear if DUI obtains with *z*-evaluative suffixed words. In fact, if DUI applies in these cases, the result is ambiguous, i.e. it is not clear whether DUI has applied or only the second coordinate element has a *z*-evaluative suffix. Thus, in the sequence *um café e um chazinho* (a coffee and a tea-DIM) it is not possible to know whether the suffix modifies both members of the coordination or only the second one. This issue is left for future research.

³⁰⁷ We may point out, furthermore, that deletion under identity in EP appears to be restricted to apply to derived words and morphological compounds. In fact, we found no instances of DUI with word+word compounds (e.g. **verde-água e verde azeitona* lit. 'green water and green olive'). With respect to V+N morphosyntactic compounds the following forms could be seen to allow DUI:

(i) *um abre-latas e abre-garrafas* > *um abre latas e garrafas*
'a bottle opener and a can opener'

um porta-óculos e porta-canetas > *um porta óculos e canetas*
'a glass holder and a pen holder'

However, these cases are probably distinct, since the interpretation here is such that the same object opens cans and bottles / holds glasses and pens. Thus, the elements that are coordinated in these structures appear to be the second members of the compound, rather than the whole compound words, as in the cases with derived words and morphological compounds. All these issues are addressed in Vigário and Frota (to appear), where a full account of DUI in Romance languages is proposed.

³⁰⁸ An interesting pair that clearly shows that the morphosyntactic properties of the original expression are kept in the clipped form is the one presented below: while the form of both clipped words is *super*, in (ia) the word is feminine (as shown by the form of the definite article that precedes it), whereas in (ib) the word is masculine.

(i) a. a (revista) *Super Interessante* > a Super the-FEM (magazine) Super Interessante

b. o supermercado > o super the-MASC supermarket

³⁰⁹ Uttered by a hairdresser.

³¹⁰ Uttered by a fireman regularly driving ambulances.

³¹¹ Uttered by Olympic Games athletes during a TV talk show.

³¹² Uttered by a nurse.

³¹³ There are a few cases of clipped words where the EP non-clipped expression does not contain two prosodic words (e.g. *metro* – cf. *metropolitano* ‘subway’; *moto* – cf. *motorizada* ‘motorcycle’). According to Machado (1977), these clipped forms were imported from French. Consequently, they say nothing about the Portuguese rule. Interestingly, unlike in the EP clipped forms based on morphological compounds, where the last vowel of the stem is usually kept low in the clipped word, in these cases the words have been integrated within the system of the language. Thus, in the word *metro*, which is masculine, the final vowel surfaces as high, whereas the word *moto*, which is feminine, has changed into *mota* (the final vowel –a corresponds to the usual ending of feminine nouns in the language). Notice, furthermore, that we are excluding from this discussion the shortenings of personal names, whose formation is often specific. For example, while reduplication is not regularly found in other areas of the Portuguese vocabulary, it may occur with proper names (e.g. *Ricardo*>*Cacá*; *Pedro*>*Pepê*; *Jorge*>*Jojó*).

³¹⁴ In the work done on other Romance Languages (see the references above), the fact that the resulting clipped words form a disyllabic foot is taken as evidence that the prosodic word in the relevant languages is subject to minimality requirements. As we have seen, the prosodic word in EP does not seem to have a minimal size. It is therefore predicted that clipped forms may be monosyllabic. This prediction is borne out, as shown by the following data: *inter-net* > *net*; *bi-sexual* > *bi*; *e-mail* > *mail*.

³¹⁵ Notice, nevertheless, that all function words with more than two syllables are stressed (e.g. *durante* ‘during’, *todavia/embora* ‘although’)

³¹⁶ See, for example, Booij (1995: 120) and Peperkamp (1997a) for similar observations in Dutch and Italian, respectively.

³¹⁷ Following the proposal of Villalva (1994) concerning the theme index and inflection, the derivational stem in nouns and adjectives may be defined as a morphological word minus the theme index and inflection (see also d’Andrade 1988: 110 for a similar definition).

³¹⁸ To be more accurate, the morphological word node corresponds in Villalva's terms to the maximal morphological projection. For ease of exposition, we use the term *morphological word* with a similar meaning.

³¹⁹ In section 4.3.3, Villalva presents a slightly different structure for *–mente* adverbs: *–mente* attaches to the morphological word node, yielding a stem to which the theme index is added, thus forming a theme; finally, inflection is added, thus creating a new morphological word node. For the present proposal, it is irrelevant whether the whole suffix plus theme index and inflection is adjoined to the morphological (word) base, or whether the suffix is added to the morphological word node and forms a new morphological base to which the theme index and inflection are added in turn. In both cases the structure will involve an inner morphological word node (containing, a derivational stem and the inflection constituent), and an outer morphological word node.

³²⁰ Recall that the future and conditional may no longer be synchronically related to the structure that gives rise to mesoclis (cf. chapter 4).

³²¹ In the EP literature, the source of word stress in function words and prefixes is in general not discussed (see section 2.3.1). With respect to prefixes, Pereira (1999) proposes that their stress follows from the fact that they form independent prosodic words. The prosodic word status of these prefixes, in turn, is seen to result from information included in their lexical entry. We assume, by contrast, that prosodic constituency is not stored in the lexical entry of individual items, but rather that it is built according to a number of generalizations and conditions (cf. section 6.4.3).

³²² This well-formedness condition may be seen to follow from principle (d) of the Strict Layer Hypothesis, as formulated in Nespor and Vogel (1986: 7), and presented in (2iv) of section 1.1 of this book.

³²³ Notice that not all compounds display a strong-weak stress pattern in Dutch (cf. Booij 1995, Visch 1999, Nespor 1999b). This pattern alternates with the weak-strong pattern, depending on the lexical category of the members of the compound, as well as on their internal structure. Under the present hypothesis, only when the compound strong-weak pattern is found is the compound word assumed to be mapped onto a compound prosodic word. In the remaining cases, the compound word is not mapped onto a prosodic compound and thus the stress pattern is assumed to depend on the ϕ -level prominence pattern or on lexical rules of stress assignment. According to this hypothesis, a prosodic distinction between compound types is predicted to emerge in Dutch, in a way parallel to EP, where some compounds are prosodized as compound prosodic words, some are grouped directly under ϕ and some are lexicalized.

³²⁴ A more comprehensive comparison between compounding in EP and other languages is beyond the scope of the present work. We remark, nonetheless, that in Dutch the typical compounds that display the compound stress correspond to the

most frequent compound type in the language, that is the nominal compound (cf. Booij 1995: 115).

³²⁵ As mentioned in chapter 1, section 1.5.4, Nespor and Ralli (1996) and Nespor (1999b) explain the prominence relations found in compound words as following from phonological phrase prominence. Nespor (1999b) supports this idea with data of a number of languages where there is a coincidence between the compound stress pattern and the phonological phrase stress pattern (as in Greek, Turkish and several Romance languages). In order to account for the stress pattern of compounds in languages like English or Dutch, further information is admitted to be required, such as the lexical category of the members of the compound.

³²⁶ Recall that the proposal that the right boundary of lexical prosodic words is not postlexically projected was driven by the prosodic behavior of pronominal clitics (cf. section 5.5).

³²⁷ Steps (iii) and (iv) seem to overlap. However, not only is (iv) necessary, as already shown in the preceding section, but (iii) is also necessary, as we will see in a few paragraphs.

³²⁸ On the arguments that support this proposal, see Villalva (1994: 6.2.2) (see also note 29 of this chapter, for a brief review).

Chapter 7

On the reduction of clitics

1. Introduction

In this chapter, we show that many alternating forms of clitic words in EP are not obtained by the general phonological rules of the language. Rather, they must either result from reduction phenomena which mainly affect specific (classes of) words or follow from allomorphy. We hope, furthermore, that the data under analysis will contribute to the understanding of reduction phenomena in general. To our knowledge, the type of reduction studied in the present chapter has never been investigated in the literature on EP.

Several types of reduction phenomena have been described to affect clitic words in various languages. The kind of phonological relation that reduced forms establish with their unreduced counterparts, together with the contextual information for the occurrence of clitic forms, constitute the basis of the typology of stressless elements presented in Zwicky (1977). As seen in chapter 5, section 4.1.1, Zwicky classifies stressless units as (i) “simple clitics”, which are reduced variants of their strong counterparts that occur in unstressed positions, and whose shape is phonologically related to the strong form; (ii) “bound words”, which are unaccented elements that show some syntactic freedom; and (iii) “special clitics”, which are unaccented words that have a special syntactic distribution and whose relation to the strong counterpart is often non-transparent.³²⁹ This classification appears not to exhaust all the possibilities, however. The data described in this chapter will show that, like simple clitics, lexically unaccented words may show further reduced realizations, especially if they occur in unstressed positions.

As mentioned in Zwicky (1977: section 7), while some reduced forms are accountable for by means of general phonological rules, some are not, and thus must be stored as allomorphs (see, also Kaisse 1985: chap.3; Berendsen 1986; Nespor 1999a: section 4). Specific rules of reduction have also been proposed in order to account for regularities within particular paradigms. This is the case of German, where according to Wiese (1996: 7.4.3) the reduced forms of clitic pronouns are systematically related to full forms by a rule that reduces the full form’s vowel to schwa. As shown by Hall’s (1999a: section 2) description of weak/strong alternations in German

function words, the same behavior is also found in other classes of items, such as personal pronouns, determiners, conjunctions, auxiliaries and prepositions.³³⁰ By contrast, Berendsen (1986) proposes that in both English and Dutch, the weak forms of monosyllabic words are allomorphs stored in the lexicon, instead of being derived from their strong counterparts. The following arguments support his claim for Dutch (Berendsen 1986: chap.3): (i) there are cases where the phonological shape of the weak form is considerably different from the strong form, and consequently no general phonological rule of the language could derive it; (ii) in some cases the weak form has a specialized meaning; (iii) in certain expressions, only weak forms are allowed. As for EP, except for personal pronouns and two cases involving Wh-words, unstressed items do not have lexically stressed counterparts (see chapter 5, section 4.3.1). Nevertheless, we will see that stressless items often show further reduced realizations. In sections 5.1 and 5.5 we discuss whether the reduced forms described in this chapter for EP ought to be analyzed as following from allomorphy or from more superficial reduction phenomena, and if so, what generalizations may be drawn for this language.

Considering English simple clitics, in Zwicky's terminology, Selkirk (1984, 1996) observes that function words in phrase final position, focalized, or uttered in isolation have a shape similar to lexical words. By contrast, in other environments they are reported to display the properties of stressless syllables. The examples in (436), taken from Selkirk (1996: 193–194), illustrate the contrast between the strong form of function words and their weak counterparts, as well as the similarity between strong/weak forms and segments in stressed/unstressed positions of full words.

(436)	Strong forms	Weak forms
	<i>for</i> [fɔːr] (cf. <i>four</i>)	<i>for Timothy</i> [fr] (cf. <i>fertility</i>)
	<i>can</i> [kæn] (cf. (<i>tin</i>) <i>can</i>)	<i>can pile</i> [kən], [kɒ], [kɪ] (cf. <i>compile</i>)
	<i>him</i> [him] (cf. <i>hymn</i>)	<i>need him</i> [ɪm], [ɪ] (cf. <i>Needham</i>)
	<i>at</i> [æt] (cf. <i>hat</i>)	<i>at home</i> [ət] (cf. <i>atone</i>)

As the data presented in chapter 5 has already shown, EP clitics do not pattern like English function words in this respect. For example, the complementizer *que* 'that' may be the only member of an intonational phrase, and still be headed by a schwa, unlike lexical words. Furthermore, clitic

function words headed by schwa or [ə] still surface with such vowels when assigned initial or emphatic stress. We have seen in chapter 3 that schwa and [ə] may not appear in (word) stressed positions. Under the analysis proposed in this book, such pattern obtains because some reduction phenomena in EP are operative within the lexical component, and thus may not be affected by postlexical phonological information. The data described in section 4 will show, in addition, that other reduced forms also occur in EP, in particular when clitics appear in unstressed prosodic configurations. This may be explained by the fact that this specific kind of reduction, similar to the English cases mentioned above and unlike EP lexical vowel reduction, occurs at a more superficial level.

Other factors seem to favor (postlexical) reduction, besides the occurrence of clitics in unstressed positions. Among those factors is word frequency. As pointed out by Selkirk (1984: 7.1.3), the very frequent English words *do* and *to* may reduce (as a consequence of destressing), while the lower frequency words *through* and *too*, with identical underlying vowels, may not. Booij (1995: 130) reports the same general tendency for very frequent words to reduce more easily in Dutch. Furthermore, highly frequent words may allow the reduction of high vowels, which otherwise are not reducible in this language (see also Van Oostendorp 2000: 145). Frequency effects are also demonstrated by the results obtained by Jurafsky, Bell, and Girand (2002) for English. In this study, the reduction of very high frequency function words with different morphosyntactic functions was investigated. For an item like *to*, which may function as the infinitive marker or as a preposition or particle, it was found that the infinitive marker is realized with greater reduction than the preposition/particle. These results correlate with the relative frequency of each morphosyntactic category in the corpus analyzed, since the former occurs more frequently than the latter. The figures are shown in (437), where the average of the duration in milliseconds for each function word is given, together with their relative frequency in the corpus (based on Table 6 of Jurafsky, Bell, and Girand 2002).³³¹

(437) Morphosyntactic category	Count	%	Average duration (ms)
Infinitive marker	543	74%	109
Preposition/particle	195	26%	138

Several other factors may play a role in English reduction. Jurafsky et al. (2001) refer the following:

- Rate of speech: faster speech rate favors reduced forms;
- Speech style: informal registers favor reduction;
- Segmental context: the form of a word may be influenced by the segmental context (e.g. consonant deletion is favored if the segment is preceded or followed by a consonant);
- Phrase boundary position: words in initial or final position may show different durational patterns;
- Contextual prominence pattern: metrical prominence may affect relative durations;
- Probability of a word given neighboring words: the conditional probability of a word given adjacent words has been shown to play a role in reduction;
- Social variables: factors such as age, gender, race, social classes have also been reported to influence reduction.

The kind of speech also constitutes an important variable to take into consideration. Keating (1998) compares the pronunciations of 40 lexical items of English in read and in non-read speech corpora, and concludes that non-read speech is less variable.³³² Bybee (2000) reports, additionally, that reduction may depend of the status of words within the discourse. This follows from the work of Fowler and Housum (1987), who found that the first use of a given word in a spoken text was longer than in subsequent uses.³³³

Vowel reduction in Dutch has also been described to be dependent on several factors. Besides word frequency, reduction appears to depend on the type of vowel involved and its prosodic and metrical position. According to Booij (1995: 6.4), high vowels and vowels in syllable initial position do not reduce easily. In addition, vowels in word final syllables do not usually reduce. By contrast, reduction is more common in interstress positions, in open syllables, and, in particular, in closed syllables in word internal position. Furthermore, some reduced forms seem to occur typically only in certain registers (see also Van Oostendorp 2000: section 2). The importance of style register to the shape of clitic words has also been pointed out for German, where, according to Wiese (1996: 250), some realizations of clitic pronouns are only found in certain registers.

One of the main goals in the following sections, is to determine (some of) the factors that may (dis)favor the occurrence of reduced forms of EP clitic words.

As this brief presentation already shows, it is not always clear whether certain reductions are specific of stressless function words, or affect instead unstressed positions in general. It was seen above that the weak forms of English function words undergo the general reduction also found in stressless positions of full words. Nevertheless, with certain English auxiliary verbs the language general tendency for reduction is considered to be carried further (cf. Selkirk 1984: 401, and, for more general discussions of auxiliary reduction in English, see Kaisse 1985: chap.3, and Nespor 1999a). With respect to European Portuguese, we have already seen that word internal vowels in stressless positions are usually subject to a process traditionally dubbed “vowel reduction” (see chapter 3, section 3). In addition to vowel reduction, other more general reduction phenomena are also found in this language. In the discussion conducted in section 21 we try to disentangle the two types of phenomena in EP. In sections 5.1 and 5.2 we will further show that some reductions are particularly evident only with specific (classes of) words.

Most forms described in this chapter are likely to be felt by the speakers of the Lisbon variety of EP as possible realizations of the relevant clitics. However, the context that (dis)favors the occurrence of one form or another is much more difficult to determine (if at all possible) if we resort only to speakers’ intuitions. Given our goal to identify some of the contexts that may affect the reduction of clitic words, we have collected external data bearing on this issue. The procedure followed for obtaining the materials, together with their description, is given in section 3. The results are presented in section 4 and discussed in section 5. In addition to establishing a few generalizations concerning the reduction phenomena described, an attempt is also made to relate our findings with other facts of EP, such as the distribution of pronominal clitics, as well as the lexicalization of sequences of function words. Finally, we touch on the issue of weighing allomorphy against reduction. Section 6 closes this chapter with a summary of our main conclusions.

2. EP vowel reduction and other reduction phenomena

Our primary task in this section is to distinguish between the process of vowel reduction (VR) described in chapter 3, section 3 and other reduction phenomena that EP also shows.

A major difference between the two kinds of reduction follows from the locus in grammar where each phenomenon is active. As shown in chapter 3, VR has the properties of a lexical rule. Although it certainly is very general, it has many exceptions, some of which dependent on morphological information. Additionally, it is usually obligatory, that is, when it may apply, it necessarily does so. Recall, nonetheless, that reduced (raised and centralized) vowels such as [i] and [ɐ] correspond to full vowels underlyingly (see, among others, Mateus 1975; Andrade 1980; Mateus and d'Andrade 2000). This is clearly shown by the pairs of words provided in chapter 3, section 3, some of which repeated in (438) below, where the same underlying vowel appears as a full vowel when it bears word stress and as a reduced vowel in unstressed position. Furthermore, [i] and [ɐ] are not found in stressed positions (except in nasal or palatal environments, in the case of [ɐ]), or in any other position where VR does not apply (see chapter 3, section 3 for further details). Their derived status is, thus, straightforward.

(438) <i>dEdo</i>	[é]	'finger'	<i>dedAda</i>	[i]	'pinch'
<i>fEsa</i>	[é]	'caress'	<i>festInha</i>	[i]	caress-DIM
<i>gAto</i>	[á]	'cat'	<i>gatInho</i>	[ɐ]	cat-DIM

Besides having exceptions and being obligatory, another indication of the lexical status of VR is the fact that it is not conditioned by factors like speech rate, or the presence/absence of non-primary stresses, which are commonly reported to affect other types of reduction phenomena (see section 1). The most natural assumption, therefore, is to consider vowel reduction a phenomenon that operates within the lexical component.³³⁴

Other reductions in EP appear to have a more general distribution. Specifically, it is well-known that schwas may usually be deleted in normal speech (e.g. Morais Barbosa 1965; Andrade 1980; Mateus and Delgado Martins 1982). The same is true of [u] in stressless position, in particular if it is word final. Indeed, as the results of a perceptual test reported in Mateus

and Delgado Martins (1982) show, the deletion of [i] and [u] often leads to the perceptual confusion of pairs of words like those in (439).

- | | |
|------------------------|--------------------------------|
| (439) <i>rape/rapo</i> | (he) shave-SUBJ/ (I) shave-IND |
| <i>decente/docente</i> | ‘decent’/‘teacher’ |
| <i>tome/tomo</i> | (he) take- SUBJ/ (I) take- IND |
| <i>apertar/aportar</i> | ‘(to) press’/ ‘to shore’ |

As for the other vowels, usually they are not deleted. With respect to [ɐ], as far as our dialect is concerned, the only context where its deletion may be accepted is in non-final posttonic positions (e.g. *Alcântara* ‘id.’, *lâmpada* ‘lamp’, *câmara* ‘chamber’), even though we find the deletion of this vowel very marked.³³⁵ Both [ɐ] and [u] can also be deleted in word final position when followed by a word starting with a vowel. However, in this case deletion is not driven by reduction, but is rather a hiatus resolution process (see chapter 3, section 14 and 13 for [ɐ] and [u], respectively).³³⁶

As for [i], it may not be deleted, unless it occurs in specific contexts where deletion follows from dissimilation or assimilation processes: for example, an [i] followed by a syllable headed by [i], or by a palatal consonant, is usually deleted (see d’Andrade 1994a, and chapter 3, section 6).³³⁷ Thus, in these cases [i] deletion is not triggered by reduction either. Finally, deletion is also impossible for the remaining unstressed vowels, which are always instances of the exceptional non-application of VR. Illustrative examples where vowel deletion may occur are given in (440a), and impossible instances of deletion are given in (440b).³³⁸

- | | | | |
|-------------------------|--------------|------------------------------------|-----------|
| (440) a. <i>devedor</i> | [dvdór] | <i>despregar</i> | [dʃprgár] |
| ‘indebted’ | | ‘to unfasten’ | |
| <i>professor</i> | [prfsór] | <i>lambu</i> | [lɐ̃b] |
| ‘teacher’ | | ‘I lick’ | |
| b. <i>opacidade</i> | [opɐsidád] | (*[pɐsidád]/*[opsidád]/*[opɐsdád]) | |
| ‘opacity’ | | | |
| <i>edital</i> | [editát] | (*[ditát]/*[edtát]) | |
| ‘edictal’ | | | |
| <i>procuração</i> | [prɔkuresẽw] | (*[prkuresẽw]/*[prɔkursẽw]) | |
| ‘procuration’ | | | |
| <i>rectidão</i> | [ɾetidẽw] | (*[ɾtidẽw]/*[ɾetdẽw]) | |
| ‘rightness’ | | | |

A number of facts show that, unlike VR, [i] and [u] deletion patterns like a postlexical phenomenon. First, it is exceptionless and its application is not obligatory (e.g. Mateus and Delgado Martins 1982). It is also a common assumption that faster speech rate and more informal styles favor it (e.g. Andrade 1980: section 5). Furthermore, it is conditioned by postlexical prosodic information. For example, the presence of a prosodic word initial (non-primary) stress disfavors the deletion of [i] and [u] in prosodic word initial position;³³⁹ similarly, these vowels tend not to be deleted in intonational phrase final position, at least when assigned a rising contour (for example, in yes-no interrogative sentences – cf. Frota 2002a). We thus conclude that this type of reduction operates postlexically.

According to the facts just presented, the hierarchy of EP stressless oral vowels deletion as a consequence of postlexical reduction is the one in (441).

(441) Hierarchy of postlexical reduction of EP oral vowels (deletion)

- a. [i] – strongly tends to be deleted;
- b. [u] – is often deleted, particularly in word final position;
- c. [ɐ] – may only be deleted (if at all) in non-final posttonic syllables;
- d. the remaining vowels – [i], [e], [ɛ], [o], [ɔ] – are usually not deleted.

We should stress, nevertheless, that our description is intended to mirror EP speaker intuitions regarding vowel deletion, but not necessarily actual productions. In fact, like in other languages further reduction, usually judged to be impossible, may be found in EP especially in informal spontaneous speech (see section 5.5).

Within the topic of the reduction that may affect clitic function words, which we describe in section 4 and discuss in section 5, it is important to further remark that stressless nasal vowels and diphthongs are not among the segments reported in EP literature to be deleted or monophthongize (e.g. Delgado Martins 1975). The deletion of [r] in the second position of a syllable onset cluster is also not a regular phenomenon of the language. More generally, consonants are not reported to be deleted, except for specific segments in particular contexts (e.g., according to Mateus 1975, among others, /l/ is deleted immediately before the nominal plural suffix³⁴⁰). These observations are supported by the acoustic study carried

out by Delgado Martins (1975), where no consonant deletion was found in a total of 364 (underlying) consonants that were observed.

Before comparing the reduction phenomena that affect stressless function words in EP with the more general reduction phenomena found within and across full words, the next section is devoted to the description of our data collection.

3. Data collection: materials and procedure

The materials that constitute the basis of our description of the phonetic shape of stressless function words consist of read texts recorded in two stages. First, we have collected materials designed to study consonant-schwa clitics when followed by a word starting with a vowel. The set of clitics studied at this stage includes complementizers, the preposition *de* ‘of’, and pronominal clitics. In a second stage, we have collected materials designed to study the remaining set of function words under observation, namely, *em* ‘in’, *ao(s)* to-the-MASC(PL), *com* ‘with’, *para* ‘for, to’ and *pelo(s)/pela(s)* by-the- MASC(PL)/FEM(PL).

In the first stage, different types of texts were manipulated so as to include as many occurrences as possible of consonant-schwa clitic items followed by words starting with a vowel. Each type of text was intended to favor the occurrence of clitic forms that might be style dependent:

- A piece extracted from the literary text “Os Maias”, written by Eça de Queirós, one of the most renowned Portuguese writers, containing 967 words, which was intended to elicit the most formal register;
- Two articles from newspapers, containing 971 words, which were intended to elicit a normal register;
- A letter written in a very familiar style, containing 1169 words, which was intended to elicit an informal register;
- A children’s tale, containing 925 words (no register was anticipated).

In the second stage, similar types of texts were manipulated so as to include as many occurrences as possible of *em*, *ao(s)*, *com*, *para* and *pelo(s)/pela(s)*:

- A piece extracted from the same literary text “Os Maias”, containing 955 words;
- Two articles from newspapers, containing 867 words;
- A letter written with a very familiar style, containing 1045 words;

- A children's tale, containing 786 words.

In order to obtain non-read material as well, we also asked speakers to retell the tale previously read.

The texts were read by three speakers in a silent room, and recorded on audiotape. The speakers were all born and have continuously lived in the Lisbon region, they have university education, and are between twenty-two and thirty two years old. In general, they present the same sociolinguistic characteristics as ourselves. All the speakers were naïve as to the purposes of the experiment.

The speakers were instructed to read each text silently, and then to read it aloud, in a natural way.³⁴¹ With the aim of evaluating the importance of speech rate in the shape of clitic function words, each text was read at two rates. First, the text was read at what the speakers considered to be a normal rate, and then it was read faster with respect to the previous reading. For the purposes of this book, the importance of the speech rate factor was only assessed for consonant-schwa clitics.

The total number of clitic words obtained for analysis is given in (442).

(442) Function words	Speakers			Total
	RI	IV	IS	
consonant-schwa clitics	439	451	454	1344
<i>em</i>	88	89	84	261
<i>ao(s)</i>	57	59	61	177
<i>com</i>	76	69	58	203
<i>para</i>	97	97	96	290
<i>pelo(s)/pela(s)</i>	49	49	49	147
Total	806	814	802	2422

We then listened to the recorded materials and transcribed phonetically each function word. The realization of prosodic words ending with an underlying non-back non-high vowel followed by words starting with a vowel was also transcribed, thus allowing the comparison between vowel drop in clitic and in prosodic word final position.

In order to identify the position of the clitic within the prosodic string, an additional task was undertaken that consisted in the identification of prosodic constituency above the word level in the actual renditions of the three speakers. However, since the perceptual identification of phonological

phrase boundaries is usually not possible in EP (e.g. Frota 1996, 2000; Vigário 1998a), only the intonational phrase breaks were marked.

4. Results

4.1. Consonant-schwa clitics

As we have seen in section 1, and as pointed out in Vigário (1998b), many factors may affect the realization of clitics. For this reason, we have carried out a preliminary observation of the corpus under study, in order to identify some of the major factors that may (dis)favor clitic reduction in EP. The clitics considered in this preliminary observation were those of the form consonant-schwa when followed by a word starting with a vowel.³⁴² The factors investigated were: (i) the quality of the following vowel; (ii) the presence/absence of word stress in the following vowel; (iii) the morphological class of the following word; (iv) speech style; (v) speech rate; (vi) the position of the clitic within the intonational phrase; (vii) the morphological class of the clitic itself; (viii) the speakers.³⁴³

The results of this preliminary observation were reported in Vigário (1998b).³⁴⁴ Of the factors considered, two yielded particularly interesting results. Specifically, speech rate and the position within the intonational phrase. As for the first factor, it is shown that clitics reduce more at the faster speech rate. As for the second factor, it is observed that clitics reduce less in intonational phrase initial position than in intonational phrase internal position.³⁴⁵ Notice that, since slower speech rates usually favor the occurrence of intonational phrase breaks (e.g. Nespor and Vogel 1986), two sets of data were taken into consideration for the evaluation of the *rate of speech* factor: one consisting of all the materials, and the other including only intonational phrase internal function words. The results of the two sets showed the same tendency for a higher frequency of reduction in the faster speech rate productions.

Vigário (1998b) also observes that there is great interspeaker variability, and that not all consonant-schwa function words show the same rate of reduction. Nevertheless, within each speaker and within each class of function words, the same sensitivity to intonational phrase position and to speech rate is obtained.³⁴⁶

Given these preliminary results, special attention is given in the following presentation to the factors *position within I*, *speech rate*, *speaker*, and *word category*.³⁴⁷

Examples of possible realizations of consonant-schwa clitic words from the collected data are provided in (443). It can be seen that consonant-schwa clitics followed by a word starting with a vowel may surface with (i) a consonant and a palatal glide, (ii) a consonant and a very reduced palatal glide or a trace of it, or (iii) just a consonant. In a few cases, the realization of these clitics with a schwa was also attested.

(443)		Speakers	Speech rate	
	Examples		normal	fast
COMPLEMENTIZER				
	“...antes de ele deixar Coimbra...”	RI	[djɛ]	[dɛ]
	before that he (to) leave Coimbra	IV	[djɛ]	[djɛ]
	‘...before he left Coimbra...’	IS	[dɛ]	[dɛ]
PREPOSITION				
	“...este tipo de assunto...”	RI	[dø]	[dø]
	‘...this kind of matter...’	IV	[dø]	[dø]
		IS	[dø]	[djə]
PROCLITIC PRONOUN				
	“...que se abrissem as janelas...”	RI	[sjə]	[sjə]
	that IMPS open-SUBJ IMP the windows	IV	[sjə]	[sə]
	‘... that they open the windows...’	IS	[sʲə]	[sə]

In the following discussion on the importance of the factors listed above for the realization of consonant-schwa clitics, we have grouped under the term *reduced* those forms produced with complete deletion of the clitic’s underlying vowel and under the term *unreduced* the complementary set that includes the forms realized without complete vowel deletion.

Let us start by considering the position within the I-phrase. The occurrence of unreduced forms as a function of the position within I is shown in the graphics of Figures 2–5 for each type of clitic (Figures 2–3 for complementizers, Figures 4–5 for the preposition *de*). We should point out that, since proclitics never occurred in intonational phrase initial position, the relevance of this factor could not be assessed for this group of words. Figures 2 and 4 plot the results obtained in the normal speech rate condition

and Figures 3 and 5 the results obtained in the fast speech rate condition, per speaker.

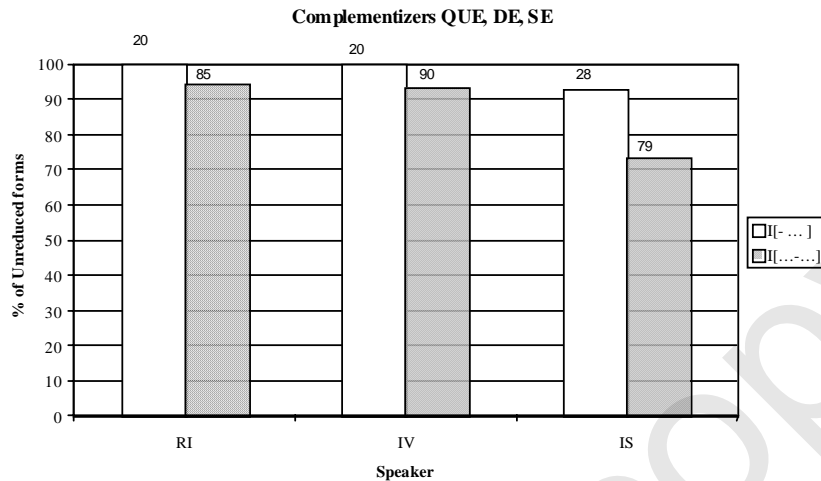


Figure 2. Percentage of unreduced forms of the complementizers *que*, *de*, and *se*, in I-initial (I[-...]) and in I-internal position (I[...-...]). Productions by the three speakers in normal speech rate. The total number of tokens obtained for each condition is given at the top of each bar.

The results plotted in the Figures 2–5 above are extremely clear. Despite the variation among speakers in the rate of reduction, the different amount of reduction depending on the type of function word, and also despite the variable magnitude of the differences related to speech rate, within each class of data the percentage of unreduced forms in I-initial position is always greater than in I-internal position. The detailed results are shown in Table 4 of Appendix II, together with their statistical analysis. The significance of the difference between unreduced and reduced forms as a function of the position in the I-phrase was tested through a one-sided significance test between two independent proportions (cf. Statistica 1997: 1460).

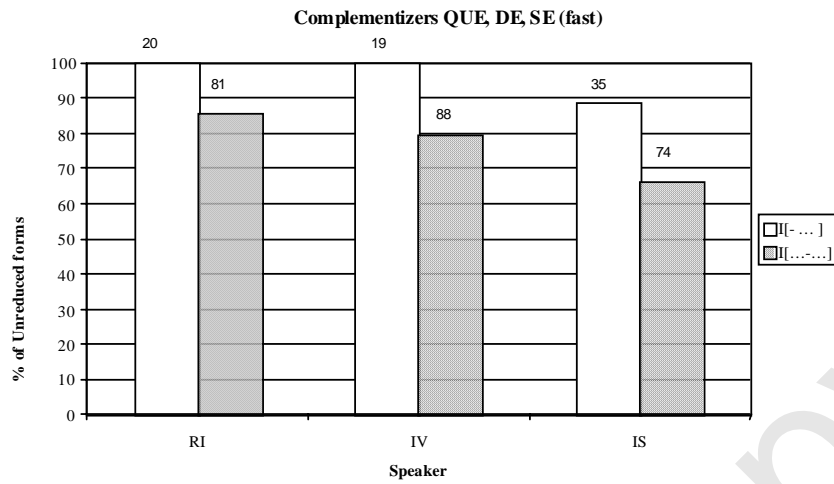


Figure 3. Percentage of unreduced forms of the complementizers *que*, *de*, and *se*, in I-initial (I[-...]) and in I-internal position (I[...-...]). Productions by the three speakers in fast speech rate. The total number of tokens obtained for each condition is given at the top of each bar.

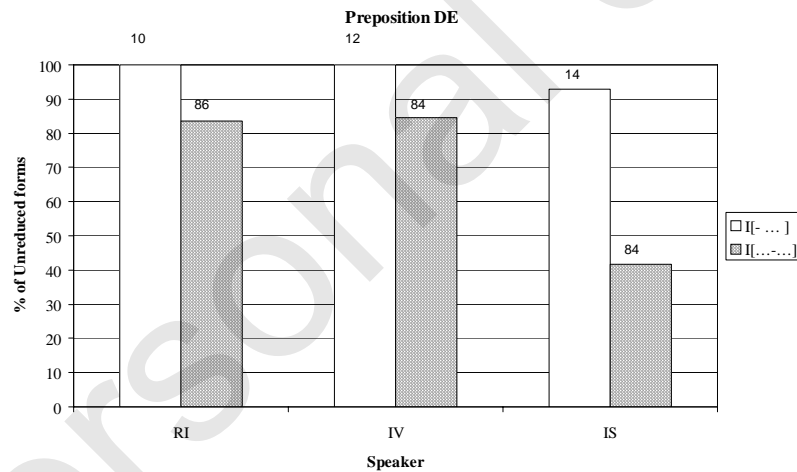


Figure 4. Percentage of unreduced forms of the preposition *de* in I-initial (I[-...]) and in I-internal position (I[...-...]). Productions by the three speakers in normal speech rate. The total number of tokens obtained for each condition is given at the top of each bar.

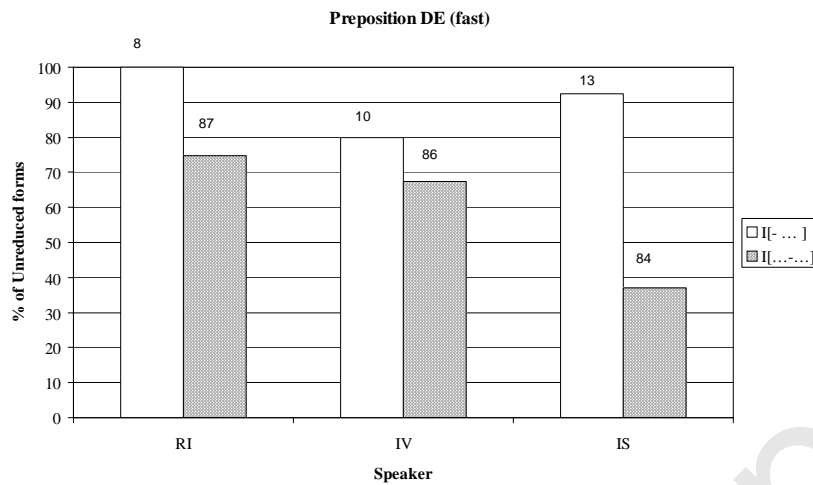


Figure 5. Percentage of unreduced forms of the preposition *de* in I-initial (I[-...]) and in I-internal position (I[...-...]). Productions by the three speakers in fast speech rate. The total number of tokens obtained for each condition is given at the top of each bar.

In the analysis performed, the results that yield $p \leq .05$ are considered statistically significant. It can be seen that the difference in the occurrence of unreduced forms in the two positions is often statistically significant. In fact, among all the factors considered, this one shows the strongest correlation with the shape of clitic words. It can thus be concluded that I-initial position clearly favours the presence of unreduced forms.

Let us now observe the remaining variables, starting with speech rate. Here, we may include the data of proclitic pronouns which only occur within the intonational phrase domain. The graphics in Figure 6 show the difference in the realization of unreduced forms in normal speech and in fast speech. The results are given separately for each speaker and for each position that clitics occupy within the intonational phrase domain.

It can be seen that, when reduced forms may occur, the faster speech rate condition favors them. In fact, when no effect of speech rate is found it is because no reduced forms appeared in a given position (which is always the intonational phrase initial position). The correlation between speech rate and reduction is nevertheless weaker than the one found for the position within the intonational phrase domain. As shown in Table 5 of Appendix II, the difference between the values obtained in normal and fast speech

rate, although in the expected direction, is often statistically non-significant. Additionally, while for the previous factor the smallest difference found is of 5.9%, and differences are in general superior to 14%, the difference between the values obtained for each speech rate may be barely above zero. For example, in the realization of the preposition *de* in I-initial position by the speaker IS, the unreduced forms occur only 0.6% more in normal speech than in fast speech. However, we should highlight the fact that no negative values were ever obtained for the difference between normal speech and fast speech results. That is, in no case there were more reduced forms in normal speech rate than in fast speech. We conclude from this that faster speech rate may, but apparently does not have to, favor the occurrence of the reduced forms of schwa-clitic words.

I-initial function words

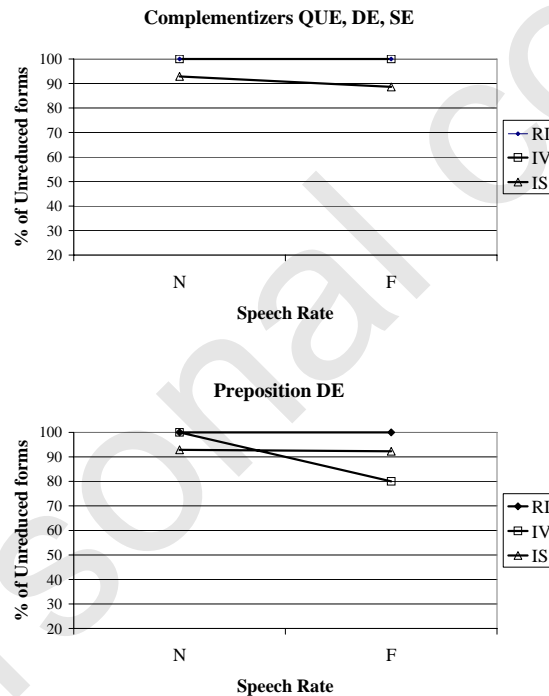


Figure 6A. Function words in I-initial position. Percentage of unreduced forms of function words (complementizers, the preposition *de*) in normal (N) and fast (F) speech rate. Results by speaker.

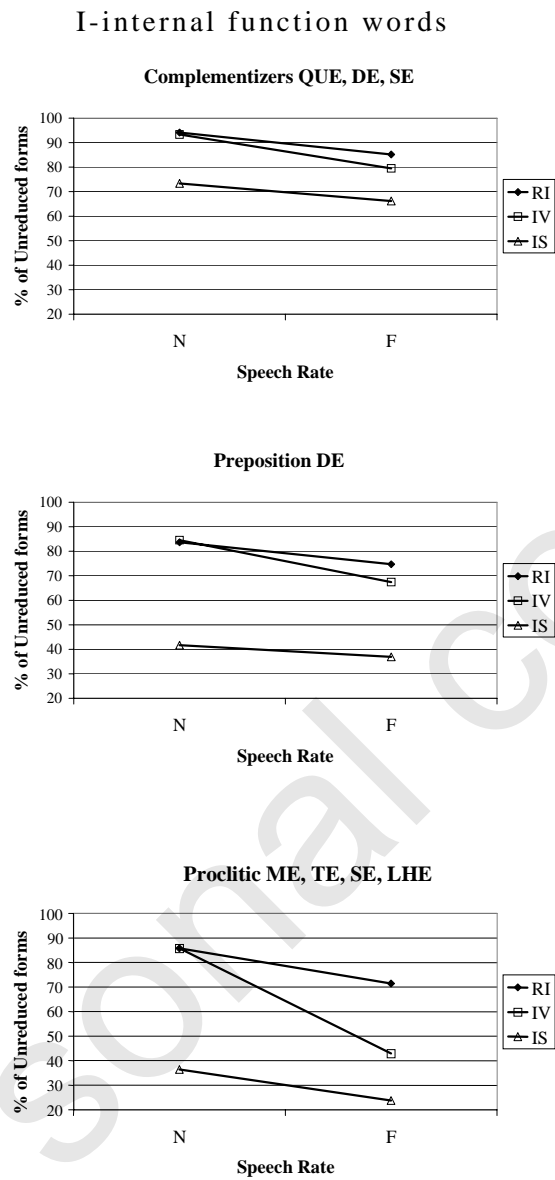
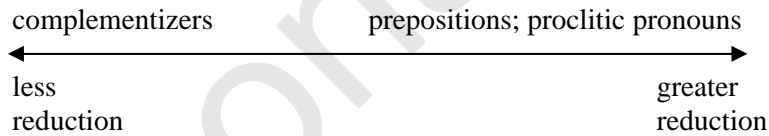


Figure 6B. Function word in I-internal position. Percentage of unreduced forms of function words (complementizers, the preposition *de*, and proclitics) in normal (N) and fast (F) speech rate. Results by speaker.

Speech rate	Speaker	Type of clitic		
		complementizer	preposition	proclitic
Normal	RI	5.9	16.3	14.3
	IV	6.7	15.5	14.3
	IS	26.6	58.3	63.6
Fast	RI	14.8	25.3	28.5
	IV	20.5	32.6	50.0
	IS	33.8	71.6	76.2

(445) Hierarchy of reduction by category



Finally, the analysis of the amount of reduction by speaker shown in (447) allows us to determine the relevance of interspeaker variation in the realization of consonant-schwa clitics.

(446) Percentage of reduced forms in I-initial position

Speech rate	Speaker	Type of clitic		
		complementizer	preposition	proclitic
Normal	RI	0	0	—
	IV	0	0	—
	IS	7.1	7.1	—
Fast	RI	0	0	—
	IV	0	20.0	—
	IS	11.4	7.7	—

(447) Speakers

	RI	IV	IS
% of reduced forms in C-schwa clitics	14.1	17.6	39.2

Of the three speakers whose data is being examined, IS presents a much greater tendency for using reduced forms, when compared to RI and IV. Consequently, we may conclude that individual differences may also yield variation in the rate of reduction.

To sum up, the shape of consonant-schwa clitics appears to be strongly influenced by their position within the intonational phrase domain, since I-initial position clearly disfavors the occurrence of reduced forms. Speech rate and type of function word also play a role in reduction: faster speech rate usually correlates with a higher percentage of reduced forms when compared to normal speech rate; and complementizers show less reduced forms than the preposition *de* and pronominal clitics. Finally, the tendency for producing reduced or unreduced clitic forms is to some extent also speaker-dependent.

4.2. Other clitics: *para*, *pelo/pela*, *ao*, *em*

Given the interest of our findings with regard to the position within the intonational phrase domain, we have decided to evaluate the importance of this factor in the reduction of the remaining group of clitics. In the case of

com ‘with’, however, other factors besides the position within the intonational phrase seem to play an important role in the occurrence of reduced forms, namely the segmental context and also the category of the following word. For this reason, we study this preposition separately, in section 4.3 below.

Examples taken from our corpus of realizations of the function words *para* ‘for’, *pelo(s)/pela(s)* for-the-MASC(PL)/FEM(PL), *ao(s)* to-the-MASC(PL), and *em* ‘in’ are given in (448).

(448) Speakers productions

Examples	RI	IV	IS
Clitic: <i>para</i>			
“ <i>Foram para França...</i> ” ‘(They) went to France...’	[prə]	[pərə]	[prə]
Clitic: <i>pelo</i>			
“ <i>...originado pelo pagamento...</i> ” ‘...caused by the payment...’	[pelu]	[plu]	[plu]
Clitic: <i>ao</i>			
“ <i>...até ao mole país...</i> ” ‘...until (to-)the smooth land...’	[ɔ]	[aw]	[ɔ]
Clitic: <i>em</i>			
“ <i>...onde em cada face...</i> ” ‘...where in each face...’	[ẽ]	[ẽj]	[ẽj]

Besides the possibilities illustrated above, intermediate forms without total vowel deletion, and consequently without resyllabification, in the case of *para* and *pelo*, and without complete monophthongization in the case of *ao(s)* and *em*, were also found, although less frequently.

Figures 7 to 10 show the percentage of unreduced forms per speaker according to the position occupied by each function word within the intonational phrase domain. Unreduced forms are assumed to be those without complete vowel deletion and resyllabification in the case of *para* and *pelo(s)/pela(s)*, and without complete monophthongization in the case of

ao(s) and *em*. We should remark that for the speaker IV, the word *pelo(s)/pela(s)* was never uttered with an unreduced form.³⁴⁸

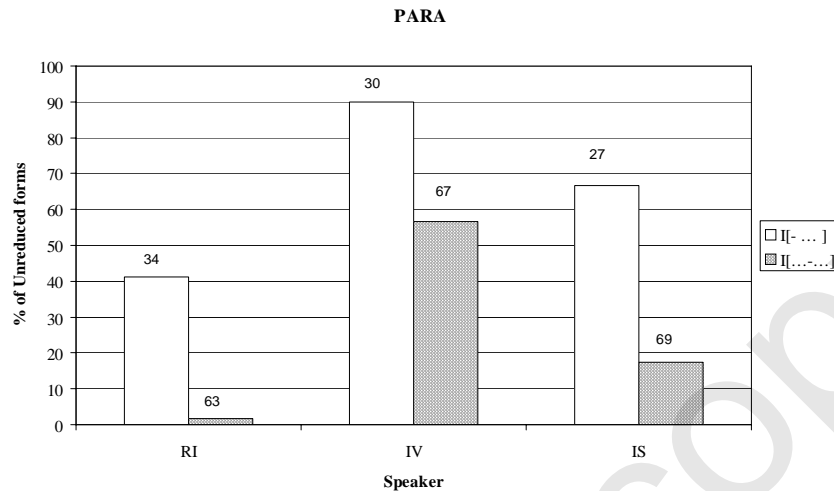


Figure 7. Percentage of unreduced forms of *para* in I-initial (I[-...]) and in I-internal position (I[...-...]). Productions by the three speakers in normal speech rate. The total number of tokens obtained for each condition is given at the top of each bar.

Again, similar to consonant-schwa clitics, the results of the reduction of the function words *para*, *pelo/pela*, *ao* and *em* with regard to the position within I are very clear. For all these clitics, the percentage of unreduced forms in intonational phrase initial position is clearly higher than in I-internal position. As shown in Table 4 of Appendix II, in the large majority of cases the difference tested is statistically significant.

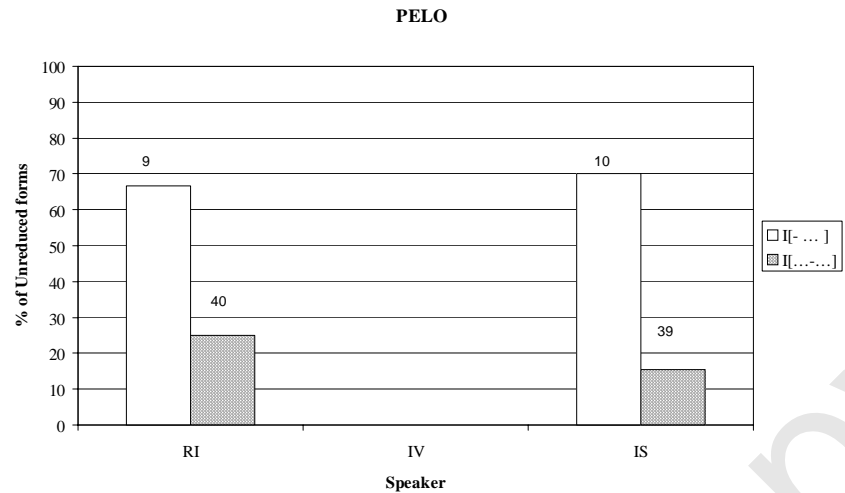


Figure 8. Percentage of unreduced forms of *pelo(s)/pela(s)* in I-initial (I[-...]) and in I-internal position (I[...-...]). Productions by two speakers in normal speech rate. The total number of tokens obtained for each condition is given at the top of each bar.

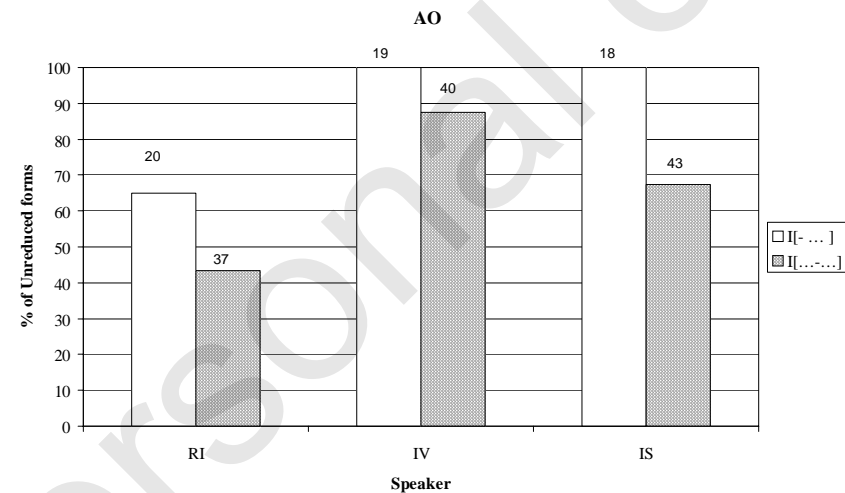


Figure 9. Percentage of unreduced forms of *ao(s)* in I-initial (I[-...]) and in I-internal position (I[...-...]). Productions by the three speakers in normal speech rate. The total number of tokens obtained for each condition is given at the top of each bar.

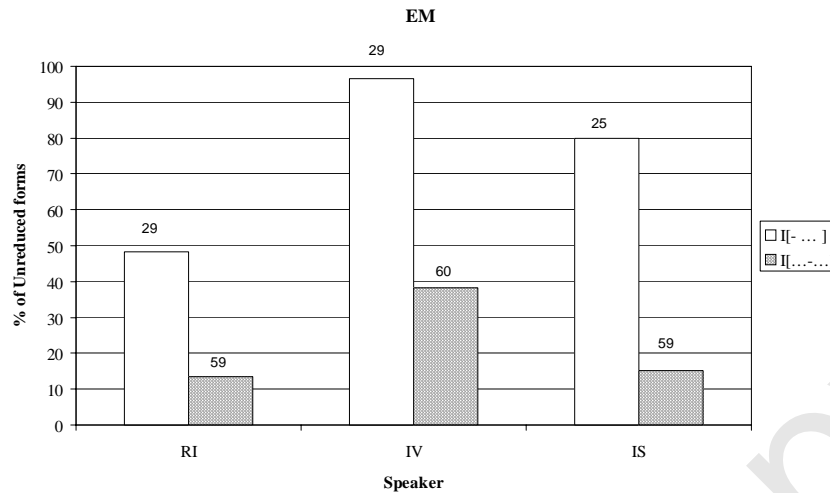


Figure 10. Percentage of unreduced forms of *em* in I-initial (I[-...]) and in I-internal position (I[...-...]). Productions by the three speakers in normal speech rate. The total number of tokens obtained for each condition is given at the top of each bar.

It should be observed, to conclude, that like what we have seen to happen with consonant-schwa clitics here there is also great variation among speakers and among clitic words in the reduction results.

4.3. The clitic *com*

Com [kõ] ‘with’ patterns differently from the clitics described in the preceding sections in that the segmental material and the grammatical category of the following word affects its realization. Once we control for these variables, however, the position within the intonational phrase domain may be shown also to constrain the shape of this function word in the expected direction.

In contrast with the disyllabic words *para* and *pelo(s)/pela(s)*, and unlike the words containing diphthongs *ao(s)* and *em*, *com* is always realized in its stronger form [kõ] when followed by a word starting with a consonant. [kõ] is also used when the following word starts with a vowel (in our corpus, these vowels are [i], [e], [ɛ], [ə], [o], [u], [ũ]). Examples of the

realization [kõ] are given in (449a). However, if the following vowel belongs to the definite article *o* ([u]) / *a* ([ø]), or to the indefinite article *um* ([ũ]) / *uma* ([umø]), besides [kõ] other realizations are also possible. When the definite article *o* follows *com*, the nasal vowel of the preposition usually denasalizes, but it may also be deleted, as illustrated in (449b).

(449) <i>Com</i> Examples	Speakers productions		
	RI	IV	IS
a.			
“...à uma hora com champanhe...”	[kõ]	[kõ]	[kõ]
‘...at one o’clock with champagne...’			
“... com horror...”	[kõ]	[kõ]	[kõ]
‘...with terror...’			
“... penteou com ele...”	[kõ]	[kõ]	[kõ]
‘...(she) combed with it...’			
b.			
“... com o rei...”	[kou]	[kou]	[kou]
‘...with the king...’			
“... com o ministro...”	[kõu]	[kõu]	[ku]
‘...with the minister...’			
c.			
“... com a relvinha...”	[kwø]	[kwø]	[kwø]
‘...with the grass-DIM’			
“ Com a queda...”	[kõø]	[kõø]	[kwø]
‘With the fall...’			
“... com a conversa...”	[koø]	[koø]	[kõø]
‘...with the chat...’			
d.			
“... com uma larga...”	[kõu]	[ku]	[ku]
‘...with a large...’			
“... com um ambiente...”	[kõ]	[kõũ]	[kõ]
‘...with an environment...’			

When the preposition is followed by the definite article *a*, its vowel usually denasalizes and semivocalizes. In a few cases, it may simply denasalize. This is illustrated in (449c). Finally, when the preposition is followed by

um or *uma*, there are several possibilities: (i) the sequence may be realized as [kõũ]/[kõumə]; (ii) the vowel of the preposition may be deleted; or (iii) the following vowel may be deleted. This is illustrated in (449d).³⁴⁹ Given these facts, we only expect to find effects of the position within the intonational phrase domain in the realization of *com* when articles follow. The results obtained for the relation between the shape of *com* and its location within the I-phrase are summarized in (450).

(450) <i>com</i> (plus article)	Forms with full vowels (nasal or oral)	Forms with semivocalization or vowel deletion
I-initial	44.9%	10.9%
I-internal	55.1%	89.1%
No. of tokens	69	46

The data in (450) allow us to conclude that I-initial position clearly disfavors the occurrence of reduced forms of *com*, that is, forms with semivocalization or vowel deletion.

4.4. Summary

The results of the production task presented in this section show a strong correlation between the position within the I-phrase of the function words studied, and the way they are realized. It was seen that, for each function word, the most reduced forms tend not to occur in intonational phrase initial position.

Besides prosodic configuration, other factors were also shown to be related to the shape of clitics, in particular of consonant-schwa clitics and *com*. As for the former set of words, we have seen that fast speech rate tends to favor the occurrence of reduced forms. In addition, not all consonant-schwa clitics show the same tendency for reduction, since complementizers appear to reduce less frequently than prepositions and pronouns. The results further show that some speakers use reduced forms more frequently than others. As for *com*, this is the only function word studied whose shape is dependent on the grammatical category of the following word, since denasalization, semivocalization, and vowel deletion occur in the great majority of cases when the preposition is followed by an article.

5. Discussion

5.1. The phonetic shape of clitics and EP general processes of reduction

The clitic function words whose form is discussed in this chapter are those we could identify so far which show reduced forms that do not pattern as expected if reduction is merely seen as a general process that applies to stressless syllables. Our goal in this section is to demonstrate this claim.

The clitic function words studied here are listed in (451), together with some of their phonetic shapes. A systematic comparison between sequences of stressless segments belonging to clitics and to full words is presented below.

(451) Clitic function words		Attested phonetic shapes
<hr/>		
C-schwa clitics + ω (V...		
e.g. <i>de</i>	‘of’	[di], [dj], [d]
<i>que</i>	‘that’	[ki], [kj], [k]
<i>para</i>	‘to’	[pæɐ], [præ], ([pæɾ]), [pæ]
<i>pelo</i>	‘by the’	[pelu], [plu]
<i>ao</i>	‘to the’	[aw], [ɔ]
<i>em</i>	‘in’	[ẽj], [ẽ], [i]
<i>com</i> + definite article/ indefinite article	‘with the/a’	[kõ], [ko], [ku], [kw], [k]
<hr/>		

As we have seen in the previous section, besides the phonetic shapes indicated in (451) each clitic may also exhibit intermediate forms. For example, the first vowel of *para* and *pelo* may reduced to the extent that it becomes almost absent, although it is still perceptually clear that there is no resyllabification. The same happens with *de* and *que*, which may be realized with a very short almost absent glide, but with no syllable restructuring, when followed by a word starting with a vowel. Similarly, *ao* and *em* may show almost total monophthongization. And the same gradual reduction is found with *com*, which may be produced for instance as a labialized plosive [k^w], when followed by the article *a*. These facts strongly suggest that the realization of these clitics is, at least to a certain extent, the result of

(gradient) reduction phenomena, rather than of lexically stored allomorphs (in section 5.5 we return to this issue).

Let us now compare in a systematic way the phonetic shape of the clitics under observation with similar segmental sequences within and across full words, starting with the clitics headed by schwa.

The realization of these clitics with a glide is only possible when the clitic is followed by a word starting with a vowel. As shown in (452a), the clitic's vowel may also be deleted. The forms in (452b), which exhibit similar underlying sequences within the word, demonstrate that this type of deletion is not a regular word internal process.

- (452) a. *gosto de animais* [djɐ]/[dɐ] '(I) like animals'
 (I) like of animals
quero que obrigues [kjo]/[ko] '(I) want you to force'
 (I) want that (you) force
- b. *adiar* [dja]/*[dã] '(to) postpone'
viajar [vjɐ]/*[vɐ] '(to) travel'

The examples in (453) show, however, that when the same vowel ends a prosodic word it is also deleted.

- (453) *perde amigos* ^x[djɐ]/[dɐ] '(he) loses friends'
balde arranjado ^x[djɐ]/[dɐ] 'pail fixed'

Nevertheless, as already seen in chapter 3, section 12, a number of aspects distinguish clitic schwa deletion from prosodic word final schwa deletion. For instance, vowel deletion in prosodic word final position is (nearly) obligatory, unlike clitic schwa deletion. Additionally, the process of schwa deletion in prosodic word final position has a categorical output, whereas schwas belonging to clitic words are often realized with a very reduced and almost absent glide, as mentioned above. The distinction between the two types of process is reflected in the data. An inspection of the realization of the words in our corpus that end with an underlying non-back non-high vowel and are followed by a word starting with a vowel shows that, unlike clitics, in these cases either the final vowel is deleted or (in very few cases) it is realized as a glide. That is, gradient realizations never occur. This contrasts with the realization of clitics, since in 11% of the productions in our corpus the clitic vowel was perceived as an almost deleted very short glide.

The data on the realization of the prosodic word final non-back non-high vowel followed by a word starting with a vowel in non-compound environments is given in (454) (where [j] stands for the realization of the glide, 0 stands for its deletion, and [j]~0 indicates an intermediate realization with a very reduced almost deleted glide).

(454) Schwa in prosodic word final position
followed by words starting with a vowel

	Normal speech rate	Fast speech rate
[j]	18	15
0	275	275
[j]~0	0	0

The results in (454) further show that prosodic word final vowel deletion is not conditioned by the rate of speech, unlike clitic schwa deletion. In addition, there is one factor responsible for the few realizations of the prosodic word final vowel as a glide. In 90% of the cases where [j] surfaces, the following vowel bears word stress.³⁵⁰ This again distinguishes this vowel deletion process from the reduction phenomenon that affects clitic words, since in the latter case, the presence of word stress in the following vowel does not seem to disfavor vowel deletion.

Another set of data that offers a clear indication that we are facing two distinct phenomena, and that further supports the analysis proposed in chapter 5 for the prosodization of clitics, is the contrast between the realization of proclitic and enclitic pronouns. Recall that we have argued for the incorporation of postverbal pronominal clitics into the prosodic word that includes the verb. Thus, enclitic final non-back vowels become prosodic word final, and therefore deletion is assumed to apply like in other prosodic word final positions. As for proclitics, they were claimed to adjoin to the following prosodic word, and thus were assumed not to undergo non-back prosodic word final deletion. The fact that proclitic vowels can nevertheless be deleted was seen to follow from an independent phenomenon of reduction (or from allomorphy), which is the topic investigated in the present chapter. The production data on enclitic and proclitic pronouns is given in (455). These results clearly set enclitics and proclitics apart. Enclitic non-back vowels behave just like prosodic word final in that they are usually deleted as a consequence of prosodic word final non-back vowel deletion, whereas proclitic non-back vowels behave like other proclitic schwa vowels that may show gradient reduction in weak prosodic configurations.

(455)	Enclitic pronouns	Proclitic pronouns
[j]	2	59
0	217	54
[j]~0	0	14

Let us now consider the remaining set of function words. We will resort to our intuitions for the description of the behavior of full words that present the same underlying sequences as the clitics under investigation.

Possible realizations of *para* found in our corpus are illustrated in (456).

- (456) *vais para casa* [pəɾɐ]/[prɐ]/[pɐ] ‘you go home’
 (you) go to home
enviado para ti [pəɾɐ]/[prɐ]/[pɐ] ‘sent for you’

These forms contrast with similar underlying segmental sequences within a prosodic word, which do not exhibit the same possibilities. As shown by the examples in (457), [ɐ] is usually not deleted within a prosodic word, although it can surface as a very short vowel. As mentioned in section 2, for some speakers of the Lisbon variety of EP, this vowel may be deleted in posttonic non-final position. However, in all varieties of EP, the deletion of a whole syllable, as in [təɾɐ]>[tɐ] or [pəɾɐ]>[pɐ], is impossible within a prosodic word.

- (457) *cítara* [təɾɐ] /??[trɐ]/*[tɐ] ‘cithara’
parallel [pəɾɐ]/?*[prɐ]/*[pɐ] ‘parallel’

This means that the forms of the clitic word *para* are not explained by a more general process of reduction that equally affects stressless syllables in EP.

With respect to the preposition+article clitic *pelo(s)/pela(s)*, it is not possible to compare it with similar sequences in unstressed positions within a prosodic word. Here, no examples were found of unstressed /e/ that did not undergo lexical vowel reduction, and therefore we cannot assess whether or not the deletion of this vowel is possible within full words. The same lack of comparable conditions makes it impossible to contrast the behavior of the preposition *com* followed by a vowel with similar unstressed sequences, within and across prosodic words. In the latter case, however, the fact that the reduced forms of *com* may only occur with articles already

constitutes an indication that this type of reduction does not follow from a general process.

Let us now consider the monophthongization of *ao(s)*, as illustrated in (458). As observed in chapter 5, monophthongization occurs in several contexts. However, the examples in (459) show that this is not a regular process that unstressed diphthongs of the form [aw] may undergo (see also the exemplification in chapter 5, section 4.3.1).

- | | | |
|-------------------------------|-----------|--------------------------|
| (458) <i>ele foi ao Porto</i> | [aw]/[ɔ] | ‘he went to Oporto’ |
| he went to-the Oporto | | |
| <i>disse ao Pedro</i> | [aw]/[ɔ] | ‘he told Peter’ |
| (he) told to-the Peter | | |
| (459) <i>exaurIdo</i> | [aw]/*[ɔ] | ‘exhausted’ |
| <i>inaudIto</i> | [aw]/*[ɔ] | ‘unheard’ |
| <i>pauIIsta</i> | [aw]/*[ɔ] | ‘inhabitant of S. Paulo’ |

As for the preposition *em*, it may also surface either with a nasal vowel and glide, or simply with a nasal vowel, as illustrated in (460). Monophthongization of nasal diphthongs, however, has not been described to be possible in prosodic word final position in the literature on the Lisbon variety of EP (cf. section 2). In addition, [ĩ] is certainly not a possible realization for the forms in (461).

- | | | |
|----------------------------|----------------|-------------------------------|
| (460) <i>está em casa</i> | [ẽj], [ẽ], [ĩ] | ‘He is at home’ |
| (he) is in home | | |
| <i>pensou em tudo</i> | [ẽj], [ẽ], [ĩ] | ‘He thought about everything’ |
| (he) thought in everything | | |
| (461) <i>bAtem</i> | [ẽj]/*[ẽ]/*[ĩ] | ‘(they) hit’ |
| <i>bEbem</i> | [ẽj]/*[ẽ]/*[ĩ] | ‘(they) drink’ |

Once again, these facts suggest that the monophthongization and raising found in the realization of the clitic *em* are not general processes that apply to unstressed diphthongs.

The data surveyed in this section leads to the conclusion that the reduction found in the clitic words under investigation does not regularly affect other unstressed positions. These reduction phenomena must therefore be

considered specific to the relevant function words. In section 5.5 we briefly address the question of whether these forms should be analyzed as following from phonetic reduction phenomena or from lexically stored allomorphy.

5.2. Reduction phenomena and word frequency

One of the factors reported in section 1 to play a major role in reduction in different languages is word frequency. With respect to the clitics under study, it is obvious that they belong to the set of the most frequent words of EP. The number of occurrences of these items, together with their ranking in the frequency list of Bacelar, Marques, and Segura da Cruz (1987) demonstrates this point (see 462).³⁵¹

(462) Clitics	Ranking	No. of occurrences	Clitics	Ranking	No. of occurrences
<i>que</i>	3	27 817	pelo/pela	76	1 248
<i>de</i>	2	33 160	ao	31	2 827
<i>se</i>	13	6 295	em	14	5 319
<i>para</i>	6	8 938	com	15	5 248

If word frequency does play a major role in reduction, we may expect to find reduction phenomena in very frequent non-clitic words as well. The following facts confirm this expectation.

Of the 500 most frequent words of EP listed in Bacelar, Marques, and Segura da Cruz (1987: 689–718), at least those in (463) clearly admit different possible realizations. The words signaled with a ‘+’ in (463) may only show the reduced realization, which is always optional, in specific circumstances. For instance, the initial fricative of the form *espera* may be deleted in the imperative form but not in the present (see the examples in 464). Similarly, the initial vowel of *embora* may be deleted if the word is used without the verb *ir* ‘(to) go’ introducing it, but not if the verb is present, as illustrated in (465).

(463) Words	Possible realizations		
<i>ainda</i>	[ɛ́idə]	[ídə]	‘still’
<i>até</i>	[ɛté]	[tɛ]	‘until’
<i>então</i>	[ɛ̃tɛ̃w̃] ([itɛ̃w̃])	[ɛ̃tɛ̃w̃], [ˈtɛ̃w̃]	‘then’
<i>mesmo</i>	[mɛ́zmu]	[mému]	‘even’
<i>sempre</i>	([sɛ̃pri]) [sɛ̃pr]	[sɛ̃p]	‘always’
<i>obrigado</i>	[obrigádu]	[brigádu]	‘thank you’
<i>estar</i> ³⁵²	[ʃtár]	[tár]	‘(to) be’
<i>espera</i> ⁽⁺⁾	[ʃpérɐ]	[pérɐ]	‘(you) wait’
<i>embora</i> ⁽⁺⁾	[ẽbórɐ]	[ˈbórɐ], [bórɐ]	‘(let’s go) away’
<i>grande</i> ⁽⁺⁾	([grɛ̃di])[grɛ̃d]	[gɛ̃dɐ]	‘great/big’
<i>pois</i> ⁽⁺⁾	[pójʃ]	[pój]	‘as’
<i>mais</i> ⁽⁺⁾	[májʃ]	[máj]	‘more’

⁽⁺⁾ This word may only reduce under specific conditions (see the comments in the text below).

- (464) a. *João, **espera** aí!* [ʃpérɐ],[pérɐ] ‘John, wait!’
 John, wait EMPHATIC-PARTICLE
 b. *O João já não **espera** mais.* [ʃpérɐ],*[pérɐ]
 the John already not wait more
 ‘John doesn’t wait anymore.’
- (465) a. ***Embora** lá!* [ẽbórɐ],[ˈbórɐ],[bórɐ] ‘Let’s go away!’
 away EMPHATIC-PARTICLE
 b. *Vamos lá **embora**!* [ẽbórɐ],*[ˈbórɐ],*[bórɐ] ‘Let’s go away!’
 (we) go EMPHATIC-PARTICLE away

As for *r*-deletion in the word *grande*, it seems to only occur when this adjective is on the non-recursive side of the modified nouns (i.e. before it), as shown by the contrast in (466).³⁵³

- (466) a. *É um **grande** companheiro!* [grɛ̃d], [gɛ̃dɐ]
 ‘(He) is a great/(big) companion.’
 b. *É um companheiro **grande**!* [grɛ̃d], *[gɛ̃dɐ]
 ‘(He) is a (*great/)big companion.’

The final fricative of the word *pois* may be deleted if it is part of the very frequent expression *pois não* ‘isn’t it’, used in negative echo questions, but it may not be deleted in similar segmental environments when *pois* functions as a conjunction, as illustrated in (467). Similarly, the consonant deletion that occurs with *mais* ‘more’ tends to happen only when the relevant item is followed by words it appears very frequently with. Examples of this sort are provided in (468).³⁵⁴

- (467) a. *Não vens, pois não?* [pój3nẽw̃], [pójnẽw̃]
 (you) not come, as not
 ‘You are not coming, are you?’
 b. *Veio, pois não queria perder nada.* [pój3nẽw̃], *[pójnẽw̃]
 (he) came, as (he) not wanted (to) miss nothing
 ‘He came, as he didn’t wanted to miss anything.’
- (468) a. *Não digas mais nada.* [máj3náðø], [májnáðø]
 (you) not say more nothing
 ‘Don’t say anything else.’
 b. *Não digas mais novidades.* [máj3nuvidádʃ], *[májnuvidádʃ]
 (you) not say more news
 ‘Don’t say more news.’

These facts show that the kinds of deletions that affect clitic function words are also found in other very frequent words of the language. Our investigation suggests that most of these deletions occur within the set of 500 most frequent words of EP listed in Bacelar, Marques, and Segura da Cruz (1987: 689–718). Apparent exceptions to this are the words *avô/avó* [əvó]/[əvó] ‘grandfather/grandmother’, which may be realized as [vó]/[vó], and which appear ranked as the 697th most frequent words of EP. Other forms, such as *papá* [pəpá] ‘daddy’ and *mamã* [məmẽ] ‘mummy’, which may be realized as [p:a] and [m:ẽ] (or [mẽ]), respectively, are not attested in the list. Additionally, the forms *obrigado/obrigada* ‘thank you’ are either absent from the list or they are counted as inflected forms of the etymologically related verb *obrigar* ‘(to) force’, which belongs to the 500 most frequent words of the language. If we examine the frequency list of inflected forms (Bacelar, Marques, and Segura da Cruz 1987: 432–688), we conclude that the forms *obrigado(s)/obrigada(s)* occurred only 65 times in the corpus, and therefore should be ranked among the 780th most frequent

words. Nevertheless, there is no doubt that all of these words are extremely frequent in EP in appropriate contexts. Indeed, the latter low figures are probably a consequence of the way the data were obtained, based on interviews. We may therefore conclude that if a word shows reduced forms (of which speakers are aware) then it must be a high frequency word in the language.

The converse is not necessarily true. That is, we cannot say that all very frequent words undoubtedly allow for similar reduced forms. Of the same frequency list referred to above, we extracted a few examples of highly frequent words where analogous deletions are felt to be impossible (see 469).³⁵⁵ Consequently, high frequency promotes reduction, but does not obligatorily enforces it.

(469) Word	Realization	Ranking in the list	
<i>assim</i>	[əsí]	*[sí]	23 ‘thus’
<i>aqui</i>	[əkí]	*[kí]	24 ‘here’
<i>depois</i>	[dipóʃ]	*[dipó]	27 ‘after’
<i>nós</i>	[nɔʃ]	*[nɔ]	37 ‘us’
<i>gostar</i>	[guʃtár]	*[gutár]	46 ‘(to) like’
<i>trabalhar</i>	[trəbɐˈlaɾ]	*[tɐbɐˈlaɾ]	78 ‘(to) work’
<i>enfim</i>	[ẽfí]	*[ẽf]	131 ‘at last’
<i>escola</i>	[kɔlɐ]	*[kɔ]	194 ‘school’

We should point out, nevertheless, that some factors may be conjectured to account for the difference among highly frequent words with regard to reduction. For example, it is likely that the existence/absence of highly frequent expressions where the relevant items occur combined with other elements may have an effect on reduction. This seems to be true for most of the items listed in (463) and (469). For instance, while an item like *até*, which may reduce, occurs in the highly frequent expressions *até já* or *até logo* (‘see you later’), an item like *assim* ‘so’, which may not reduce, does not occur in such highly frequent expressions.

With regard to the reduction of stressed words, we should further notice the following. In general, within stressed words, no deletion affects the nucleus of stressed syllables. In a few cases, however, this generalization seems not to hold. For instance, the stressed syllable of *nunca* [núkɐ] ‘never’ may be reduced to the prenasalization of the following plosive, or

even be completely deleted, in a sequence like *nunca mais* [ˈkɐmájʃ]/[kɐmájʃ] ‘never again’. A similar situation happens with *muito* [mũjtɐ] ‘a lot’, where the initial stressed syllable may be reduced to the prenasalization of the following [t] in the sequence *muito obrigado* [ˈtoʔbrigádu] ‘thanks a lot’; and with the word *isto* [iʃtu] (this-NON-ANIMATE), which in some dialects may exhibit a reduced form in the sequence *isto é* [ʃtwɛ] ‘that is’. Importantly, none of these reduced variants may occur in isolation. We may thus admit that the reduced variant is possible when the word has been destressed.³⁵⁶ These cases would therefore be similar to many English function words, which occur in the strong form in isolation and in the weak form in unstressed positions (e.g. Selkirk 1984, 1996, and section 1 above). Indeed, destressing may be seen as a particular case of reduction, which may thus consist of the deletion of segmental, autosegmental and/or suprasegmental material (see also section 5.3).

In conclusion, the reduction phenomena found within clitic words may be seen as instances of more general phenomena of reduction that are also apparent in other very frequent words of EP.

To sum up: (i) reduction phenomena affect EP high frequency words; (ii) not all highly frequent words show the same potential to reduce; (iii) destressing may be seen as a particular instance of reduction; and (iv) the same type of reduction found in clitics is also patent in full words.

5.3. Generalizations on the reduction of very frequent words

As we have seen above, reduction phenomena seem to be specific of only a very small set of highly frequent words. These words may undergo autosegmental and segmental deletions, semivocalization, monophthongization, and destressing. However, not all very frequent words show the same tendency to reduce. This has already been shown for full words in the preceding section. Similar to full words, not all clitics allow the same type of reduction. For example, the clitic *com* may undergo denasalization, but *em* may not surface as [e]; and although the cluster in [prɐ] (<*para*) may further reduce to [pɐ], the cluster in [plu] or [plɐ] (<*pelo/pela*) may not reduce to [pu] or [pɐ]. In addition to this, the same type of phenomenon often appears to be sensitive to different contextual information. For example, schwa may be deleted if followed by a vowel belonging to any kind of word, but the vowel of *com* may only be deleted if the following vowel is part of

a definite or an indefinite article; vowel deletion in the words *para* and *pelo/pela* may only affect the first, but not to the second vowel. These facts clearly suggest that the phenomena under investigation may not be accounted for by a general rule of reduction.

Another major difficulty for the formalization of such a rule concerns the nature of reduction. As we have seen, the phenomena described are to a certain extent gradient, and thus any account presupposing a categorical output misses their nature.

There are, nevertheless, a few features common to all reduced forms. A number of generalizations presented below concern the units targeted by reduction, whereas others relate to its result.

As seen in the preceding section, the items that exhibit reduced forms constitute a subset of the set of highly frequent words of the language. Additionally, within this class of words, in the large majority of the cases described in this chapter reduction affects the phonological material of word initial syllables.³⁵⁷ We have seen, furthermore, that reduction may consist of the deletion of (auto)segmental features (e.g. [consonantal] or [nasal]), suprasegmental features (e.g. word stress), as well as of whole segments (e.g. vowels). Nevertheless, and regardless of word size, in all cases reduction appears to affect only one syllable per word.

(470)	Examples	Syllable structure
Avoid complex onsets		
non-back glide deletion	<i>de arte</i> [djárt]>[dárt]	CGV>CV
back glide deletion	<i>com a</i> [kwɐ]>[kɐ]	CGV>CV
<i>r</i> -deletion	<i>para</i> ([pɐrɐ]>)[pɐ]>[pɐ]	CCV>CV
	<i>grande</i> [grɛ̃]>[gɛ̃]	CCV>CV
Avoid complex nuclei		
monophthongization	<i>ao</i> [aw]>[ɔ]	VG>V
	<i>em</i> [ẽj]>[ẽ]	VG>V
Avoid complex rhymes		
consonant deletion	<i>mesmo</i> [méʒ]>[mé]	CVC>CV
	<i>pois</i> [pójʃ]>[pój]	CVGC>CVG
Avoid onsetless syllables		
deletion of initial vowels	<i>até</i> [ɐtɛ̃]>[tɛ̃]	VCV>CV
	<i>avô</i> [ɐvó]>[vó]	VCV>CV

As for the result of reduction, many reduced forms exhibit sequences that are universally less marked than the unreduced counterparts. As shown in (470), many reductions result in the avoidance of marked syllable structures, such as those with complex onsets, complex nuclei, complex rhymes, and unfilled onsets. Furthermore, the deletion of the initial vowels of disyllabic clitics, as in *para* ‘for’ [pəɾə]>[pɾə] and *pelo(s)/pela(s)* by the-MASC/FEM [pelu][pele]> [plu][ple], results in the avoidance of disyllabic clitics.³⁵⁸ The actual weight of these generalizations in the account of reduction phenomena deserves, nonetheless, further investigation.

5.4. Factors (dis)favoring reduction

Among the factors identified to constrain the shape of clitic words, two seem particularly important, namely the intonational phrase initial position, and the grammatical category of the words involved. In the following sections we present some additional facts of EP that may bear some connection with these findings.

5.4.1. The importance of intonational phrase initial position

Our results on clitic reduction strongly suggest that the intonational phrase initial position has a special status. As mentioned before (see note 345), Frota (2000: 5.2.3.1) also finds a strong correlation between the form of the function word *aos* and its position within the intonational phrase domain. The correlation between the shape of clitic words and their position within the I-phrase has led us to suggest previously that the I-initial position is a marked position (Vigário 1998b: 372). Indeed, the hypothesis that the I-initial position is a prominent one may straightforwardly explain why clitics tend to appear in their unreduced form in such a prosodic configuration. In the following paragraphs, we present additional evidence in favor of I-initial prominence, and some consequences of this property in other areas of the behavior of clitics.

Independent evidence for the hypothesis that the intonational phrase initial position is prominent comes from the data presented in Frota and Vigário (2000). In this study, already referred to in chapter 5, two subjects have identified the non-primary word stresses occurring in 40 sentences of EP (20 sentences read by two EP speakers). As we have seen (chapter 5,

section 4.3.2), non-primary prominences were only found in prosodic word initial position. Furthermore, the results show that the intonational phrase initial position strongly favors the occurrence of non-primary stresses. The relevant data is shown in (471) below.

(471)	Perceived non-primary stresses	No. of prosodic word initial positions ³⁵⁹
I-initially	27 (67.5%)	40
I-internally	24 (20.7%)	116

The data on perceived prominence, together with the facts concerning the role played by I-initial position in the occurrence of unreduced forms of clitics, have led us to further hypothesize that I-initial prominence may also have some connection with the distribution of pronominal clitics in EP (cf. Vigário 1999c; Vigário and Frota 2000). The following discussion develops this idea.

It is well known that pronominal clitics in EP present a general distribution quite distinct from other Romance languages. While in EP pronominal clitics are usually postverbal, occurring in preverbal position only when preceded within a certain domain by a limited set of words, in other Romance languages pronominal clitics are usually preverbal, occurring in postverbal position when the verb appears in specific forms (as the imperative or the infinitive) (see, for example, Galves 1992; Martins 1994; Barbosa 1996; Rouveret 1999; Duarte and Matos 2000). One of the long-standing questions concerning the peculiar distribution of pronominal clitics in EP is its origin (e.g. Said Ali 1964; Brandão de Carvalho 1989; Frota 1994; Galves and Galves 1995). The fact that the Brazilian variety of Portuguese patterns like other Romance languages and unlike EP renders the question even more interesting.

In several studies, it has been proposed that the distinction between the two varieties of Portuguese that has led to differences in clitic placement is a prosodic one. Specifically, in Modern European Portuguese pronominal clitics have been hypothesized to be phonologically enclitic, which implies that they always have to be preceded by a prosodic word (cf. Brandão de Carvalho 1989, and Galves and Galves 1995, following an idea previously suggested in Said Ali 1964). BP clitics, by contrast, are assumed to be phonologically proclitic, which is compatible with their typical location before the verbal host. This analysis is capable of accounting for the impossibility

of sentence initial clitic pronouns in EP, in contrast with BP. There is, however, one major problem with such a view. As we believe to have shown in chapter 5, preverbal pronominal clitics behave like other phonological proclitics, and EP has a general tendency for proclisis instead of enclisis. Consequently, the reason why pronominal clitics may not appear sentence initially should not be ascribed to the leftward direction of pronominal cliticization.

In contrast with the proposals mentioned above, our hypothesis is that I-initial prominence may have played a role in the emergence of the enclitic pattern presently found in EP. To understand this it is important to bear in mind a few historical facts. Essentially, if we exclude the particular contexts where proclisis is still triggered in present-day EP – where proclisis is systematically found at least since the 13th century (Martins 1994) – both in Medieval and in Classical Portuguese enclisis and proclisis freely coexisted (Martins 1994; Britto 1998).³⁶⁰ This pattern is lost in favor of a generalized enclisis at some point around the 17th century (Martins 1994).³⁶¹ In order to explain this shift, we adopt the general idea of Galves and Galves (1995) that a rhythmic change occurred during this period. Contrary to these authors, however, we speculate that the modern property of EP that has emerged at that point and that played a crucial role in pronominal clitic location was the prominent status of the intonational phrase initial position. Indeed, the emergence of an initially prominent I-domain may explain the preference for enclisis in those contexts where both preverbal and postverbal locations were possible. Since both clitic locations were allowed, the choice of the postverbal pattern would always prevent the clitic from appearing in this position, where it could become stressed. Notice that postverbal pronominal clitics are described at least since Gonçalves Viana (1883) as not bearing any kind of secondary prominence in the European variety of Portuguese, and the same is true for BP. Thus, our default assumption is that at the period when the change occurred, clitics would never bear non-primary stresses in postverbal position as well. As pointed out by Frota (1994), it is interesting to observe, by contrast, that the rhythmic properties of other Romance languages, like French, and some varieties of Italian and Spanish, which exhibit a proclitic pattern, unlike EP, are such that postverbal clitics may bear secondary or even primary word stress.

In short, it is conceivable that the selection of the enclitic pattern in EP may have constituted a strategy to prevent pronominal clitics from occurring in prominent positions.

The idea that prominence factors may explain the distribution of clitics was previously suggested in Adams (1989) for other languages. One such

language is Old French, which the author speculates to have had both initial and final stresses at the word and phrase levels. The ban of sentence initial clitics in this language is explained precisely by the existence of a sentence initial stress. Adams further suggests that there is a correlation between the weakening of the initial stress and the loss of enclisis that began in Old French and continued in Middle French.³⁶² Other languages are also reported to exhibit different strategies to prevent clitics from appearing in an initially stressed position, such as German and Frisian. In these languages, clitics do not occur initially because either full pronominal forms, instead of clitic pronouns, are used in this position, or full (stressed) words, such as adverbs, verbs or nouns, are placed in initial position, thus preceding the clitic.

We should emphasize that our hypothesis for the shift of Portuguese pronominal clitics towards enclisis is not that the strong I-initial position has forced enclisis, but rather that, since both proclisis and enclisis were available at the period of the change, the choice of enclisis was preferred because this clitic placement prevented them from becoming stressed. This view explains why other proclitics, such as prepositions, articles and conjunctions have always been admitted in sentence initial position, since in these cases no alternative strategies exist in the language that would allow these items to escape from appearing in a strong position.

In other languages too, different clitics do not show equal sensitivity to initial position. For example, according to Rudin et al. (1999), Bulgarian pronominal and auxiliary clitics are restricted from being sentence initial, whereas other clitics are not. This is illustrated in (472) and (473), where capital letters signal word stress. In (472a), pronominal clitics may be preverbal because *AZ* precedes them, and thus clitics do not appear in sentence initial position. In (472b), by contrast, nothing precedes these clitics, and thus prosodic inversion, in the sense of Halpern (1995), has to apply, causing the clitics and the following prosodic word to invert their positions, as shown in (472c). Finally, the example in (473), cited from Hauge (1976), shows that other clitics, such as the phonological clitic conjunction *i*, are not prevented from occurring in initial position.

- (472) a. *AZ ti ja DAdox*
 I-NOM you-DAT-SG it-ACC gave-1PSG
 b. **ti ja DAdox*
 c. *DAdox ti ja*
 gave-1PSG you-DAT-SG it-ACC
 'I gave it to you.'

(473) *iVAN VCĖra se oBAđi, VÂRna mu gi, i si oTide*

Ivan yesterday REFL-ACC called-3PSG returned-3PSG him-DAT them-ACC and REFL-DAT went-3PSG

‘Ivan called yesterday, returned them to him, and went home.’

We should add that prosodic inversion in the Bulgarian data described above may not be explained by the directionality of pronominal and auxiliary cliticization. The fact that these clitics may occur in quasi-initial position, preceded by other clitics that are not restricted from being sentence initial clearly shows that these clitics do not require a prosodic word to their left. This is the case of the clitic pronoun *si* in (473) above, which may occur in quasi-initial position preceded by the clitic *i*. In fact, Bulgarian possesses one such directional clitic – the question particle *li* (cf. Halpern 1995: 2.5.1; Rudin et al. 1999). As expected given this property, this particle may not appear sentence initially or preceded by another clitic precisely because it must cliticize to the preceding prosodic word.

In short, Bulgarian also possesses a strategy that prevents some (but not all) clitics from being sentence initial. Although the reason for this is not given in Rudin et al. (1999), we may extend the hypothesis put forward for Modern European Portuguese and Old French to Bulgarian. That is, it is possible that in this language too sentence (intonational phrase) initial position is prominent, and thus a process only available to some clitics applies in order to prevent them from occurring in such a prominent location.³⁶³

If there is indeed a correlation between the prominence phenomena associated to the intonational phrase domain and the emergence of enclisis in EP, then we may expect not to find such I-initial strength phenomena in the Brazilian variety of Portuguese, where pronominal clitics may occur preverbally, even in sentence initial position. The data presented in Frota and Vigário (2000) on the BP variety indeed point in this direction. The same 20 sentences recorded for EP were also produced by two native speakers of BP. The sentences were then listened by two BP native speakers, who marked the location of perceived non-primary stresses. The number of perceived non-primary prominences in intonational phrase initial position in the BP and in the EP data is given in (474). The percentage of intonational phrase initial stresses in BP is reduced to about half the percentage found in EP. This fact becomes even more significant when we observe the total number of non-primary prominences perceived in the two varieties – 98 for BP versus 51 for EP. That is, BP shows nearly twice as much perceived prominences as EP.³⁶⁴

(474) I-initial stresses

Portuguese variety	Perceived non-primary stresses	No. of prosodic word initial positions
BP	13 (35.1%)	37 ³⁶⁵
EP	27 (67.5%)	40

In short, the initial position of the intonational phrase in BP does not pattern like a strong position. This is the expected result according to our hypothesis concerning the relation between pronominal clitic location and the prominence features associated to I-initial position.³⁶⁶

To sum up, we have proposed that clitics tend to reduce less frequently in intonational phrase initial position as a consequence of the prominence features associated to this position. Besides our results on the correlation between the I-initial position and the occurrence of unreduced clitic forms, we added independent evidence for these properties based on the results of a perception task described in Frota and Vigário (2000). These results show a significantly higher percentage of perceived non-primary stresses at the beginning of the intonational phrase domain, when compared to the amount of I-internal non-primary stresses. The prominence properties of I-initial position were then hypothesized to have been at the origin of the shift in pronominal clitic positioning that occurred in this language around the 17th century. This view was strengthened by the existence of other languages where clitics exhibit similar sensitivity to I-initial position. Finally, the fact that Brazilian Portuguese shows much less perceived stresses in I-initial position than EP was seen to be compatible with our hypothesis on the relation between I-initial stress and the distribution of clitics, since pronominal clitics in BP are not prevented from appearing sentence initially.

5.4.2. *The importance of grammatical category*

Our investigation of the factors that may affect the realization of clitic words has also revealed the importance of the grammatical category of clitic words, and in some cases of the following words as well. This issue is addressed in the following paragraphs.

We have seen that the preposition *com* [kõ] ‘with’ may be realized as [ko], [ku], [kw] or [k] when followed by the definite or the indefinite arti-

cle *o(s)/a(s)* and *un(s)/uma(s)*. As shown below, similar phenomena have existed in other periods of the language, causing the fusion and subsequent lexicalization of two adjacent function words.

Although the realization of the preposition *de* 'of' may usually alternate between [dj] and [d] when followed by a word starting with a vowel, when it occurs with a small set of function words the form with the glide is no longer possible. These cases are listed in (475).

- (475) a. *do, da* (<*de o, de a*)
of the-MASC, of the-FEM
b. *desse, dessa, disto* (<*de esse, de essa, de isso*)
of that-MASC, of that-FEM, of that-NON-ANIMATE
c. *deste, desta, disto* (<*de este, de esta, de isto*)
of this-MASC, of this-FEM, of this-NON-ANIMATE
d. *daquele, daquela, daquilo* (<*de aquele, de aquela, de aquilo*)
of that-MASC, of that-FEM, of that-NON-ANIMATE
e. *dele, dela* (<*de ele, de ela*)
'of him', 'of her'
f. *daqui, dali, daí, dalém, dacolá* (<*de aqui, de ali, de aí, de além, de acolá*)
'from here', 'from there', 'from there', 'from there', 'from there'

The fact that the unreduced form is no longer possible in these cases indicates that the preposition and the following function word have lexicalized as a single unit (see, for example, Napoli and Nevis 1987 for similar cases in Italian, and the discussion in chapter 4, section 4.2.2 of this book on clitic clusters in EP).

Lexicalization has also occurred with the preposition *em* ‘in’. This case is a bit more complex because the lexicalized forms are remnants of a former stage of the language when the final nasal consonant could surface at the onset of the following vowel initial syllable, across words (see Williams 1975: §143). As mentioned in chapter 3 section 5, this behavior is presently confined to derived words. The lexicalized forms are shown in (476).

- (476) a. *no, na* (*<em o, em a*)
in the-MASC, in the-FEM
b. *nesse, nessa, nisso* (*<em esse, em essa, em isso*)
in that-MASC, in that-FEM, in that-NON-ANIMATE

316 *On the reduction of clitics*

- c. *neste, nesta, nisto* (*<em este, em esta, em isto*)
in this-MASC, in this-FEM, in this-NON-ANIMATE
- d. *naquele, naquela, naquilo* (*<em aquele, em aquela, em aquilo*)
in that-MASC, in that-FEM, in that-NON-ANIMATE
- e. *nele, nela* (*<em ele, em ela*)
'in him', 'in her'

Notice that, like in the preceding case with the preposition *de*, the vowel of *em* is obligatorily deleted.³⁶⁷

Other similar cases of lexicalized forms occur with the preposition *por* (*<per*) 'by' when followed by the definite article, as illustrated in (477)³⁶⁸

- (477) *pelo, pela*
by the-MASC, by the-FEM

Finally, the preposition *a* 'to' when followed by a limited number of function words has also undergone lexicalization, as shown in (478).

- (478) *ao, à* (*<a o, a a*)
to the-MASC, to the-FEM
àquele, àquela, àquilo (*<a aquele, a aquela, a aquilo*)
to that-MASC, to that-FEM, to that-NON-ANIMATE

There are two interesting generalizations concerning the first members of these lexicalized sequences: (i) the first member is always a preposition; (ii) the set of prepositions that have undergone lexicalization is a subset of the set of prepositions that still have reduced variants in present-day EP (see section 4).³⁶⁹ In fact, of the prepositions studied, only *com* and *para* do not occur lexically combined with some adjacent material.³⁷⁰ In the case of *para*, this may be due to the fact that this is a disyllabic word. As for *com*, lexicalization has probably not occurred, since, when reduction is possible, both reduced and unreduced variants may occur. Nevertheless, this item may appear reduced only with a small set of function words, which is a subset of the set of function words that participates in lexicalized expressions with other prepositions.

The relevance of word category (generalization i) is shown by the fact that complementizers never occur in lexicalized expressions of this type even if they have precisely the same shape as prepositions. The contrasts in (479) and (480) illustrate this point.

- (479) a. *A ideia **de** o encontrar deixou-me contente.* ([du]/[dju])
 the idea of him (to) meet left me happy
 ‘The idea of meeting him has made me happy.’
 b. *A ideia **do** João era sair.*
 the idea of-the John was (to) leave
 ‘John’s idea was to leave.’
- (480) a. *Pensei **em** o convidar.* (*no)
 (I) thought of him invite
 ‘I thought of inviting him.’
 b. *Pensei **no** João hoje.*
 (I) though of-the John today
 ‘I thought of John today.’

That only the prepositions allowing for reduction may undergo lexicalization (generalization ii) is shown by the comparison between the behavior of the preposition *em* ‘in’ and the segmentally very similar preposition *sem* ‘without’: only the former may reduce, and only the former may appear in lexicalized expressions.³⁷¹

As for the second member of the lexicalized expressions, it always belongs to closed morphological classes that include very few items: articles, demonstratives, locative adverbs, and personal pronouns. Two plausible reasons for the fact that only these elements have undergone lexicalization are (i) their small number, and (ii) the fact that, unlike other classes of words, which may occur with most word categories of the language (e.g. complementizers), the combination of prepositions with the words belonging to these particular classes is very frequent.³⁷²

Besides lexicalization, clitics belonging to distinct morphological categories were also found in section 4 to behave differently with respect to reduction. Our results have shown that consonant-schwa clitics reduce significantly less in items that function as complementizers when compared to items that function as prepositions.³⁷³ The question we should address at this point is what may explain this difference.

In order to account for the difference in the frequency of reduction between complementizers and prepositions, we could hypothesize that the tendency for reduction is correlated with the frequency of words in the language (see, for example, the results of Jurafsky, Bell, and Girand 2002, presented in section 1). However, an inspection of the number of occurrences of *que* and *de* (which are typical examples of complementizer and preposition in EP) in the frequency corpus presented in Bacelar,

Marques, and Segura da Cruz (1987) shows that both types of words are extremely frequent. Furthermore, in our own corpus, complementizers occur slightly more often than the preposition *de*. The figures are shown in (481).³⁷⁴ Notice that, according to the hypothesis under test, it would be expected that the complementizer should be less frequent than the preposition, contrary to fact.³⁷⁵

(481) Function word	Total no. of occurrences in Bacelar, Marques and Segura da Cruz (1987)	Total no. of occurrences in our corpus
<i>que</i>	27 817	639
<i>de</i>	19 987	578

In conclusion, the relative frequency of these items does not seem to explain the different amount of reduction found in each class.

Besides frequency, another factor that is likely to bear on the tendency to reduction is stress. It may be expected that monosyllabic stressed items reduce less than monosyllabic unstressed ones.

In the case of *que* and *de*, it is entirely clear that both items are not assigned word-level stress. This is shown by the fact that in both cases their only vowel is a schwa, which may either surface as a glide when followed by a word starting with a vowel or be deleted. If we consider other levels of prominence, however, important differences between the two types of items emerge (see also Frota and Vigário 1996; Vigário and Frota 1998). Besides the prominence phenomena found at the word-level (i.e. the optional initial non-primary stress, and word stress), two other positions are clearly associated with prominent features in EP, namely, the intonational phrase initial and final positions. Syllables in the former position are frequently assigned a non-primary stress (see the discussion in section 5.4.1), while items in the latter position bear I-level stress and are assigned I-nuclear pitch accent (see chapter 3, section 18). Because they introduce a sentence, complementizers, unlike prepositions, often appear in I-initial position. They often occur in I-final position as well, where they are the prominent element of their intonational phrase (see chapter 5, section 4.3.1). Prepositions, by contrast, are totally excluded from I-final position in EP. Furthermore, and although they may begin intonational phrases created due to phonological factors such as balance and length (see Frota 2000), or due to the introduction of parenthetical expressions before prepositional phrases, prepositions typically

occur in I-internal positions. In conclusion, the prosodic distribution of complementizers and prepositions is such that complementizers often occur in prominent positions at the level of the intonational phrase domain, while prepositions usually do not.

Although the prominent properties associated to the positions that complementizers occupy constitute a plausible explanation for these items to reduce less than prepositions, this explanation is not sufficient. Indeed, the lower tendency for reduction is also evident when complementizers occur in I-internal (non-final) position (see Figures 2–5 of section 4.1). Given this state of affairs, we suggest that the difference between the two classes of items is possibly due to an intrinsic property of complementizers to reduce less.³⁷⁶ This property, in turn, can be the consequence of a frequency effect. Given that complementizers often occur at the edges of the intonational phrase, which are positions associated with prominent features in EP, complementizers in these positions tend not to reduce. This tendency for less reduction may have become a general property associated with this class of words, regardless of actually appearing in stressed positions or not.

Jurafsky et al. (2001) show that more probable words are more likely to reduce, and suggest that the probabilistic relations between words play a role in speakers' mental representation of language. It is possible that a similar view could explain the EP data just presented, if the information concerning the frequency with which complementizers and prepositions occur in positions where they may bear stress, and/or the frequency with which they actually bear stress, enters in the computation of reduction probability.

5.5. *Reduction versus allomorphy*

We conclude the discussion section with the delicate issue of weighing reduction against allomorphy in the account of reduced variants of very frequent words.

The main difficulty that we encounter in this domain is that reduction seems to be very superficial and phonetically driven, although it also appears to exhibit some properties of lexical phenomena. On the lexical side is the fact that many reduced forms of very frequent words are not accountable for by means of the language's rules. On the superficial side are the following facts: (i) many items show intermediate realizations, thus suggesting gradient reduction; (ii) speakers are not always aware of the possible reduced forms; and (iii) the occurrence of reduced forms is, to a

certain extent, conditioned by non-lexical information such as position within the intonational phrase domain and speech rate.³⁷⁷

Our understanding of these facts is as follows. Reduction is a general phenomenon found in the languages of the world that affects in particular very high frequency words (see section 1 for some references pointing in this direction). In some cases, reduced forms lexicalize, probably as a consequence of their very high frequency. Therefore, coexisting variants of very frequent words may either be the result of reduction of strong forms or follow from lexically stored allomorphy. Importantly, the existence of lexically stored allomorphs does not preclude the occurrence of reduction affecting the strong forms of lexical allomorphs. This view allows us not only to explain why certain types of reduction are only judged to be possible in the realization of specific items, but also to account for the gradient realizations attested. The former effect follows from a lexical feature, since only certain forms are lexicalized, thus becoming part of the explicit linguistic knowledge of speakers; whereas the latter effect is the result of reduction, which is a general phenomenon that is constantly active in the language, and of which speakers are usually not aware.

It is not always clear whether or not reduced forms have lexicalized, in particular because it can be shown that lexicalized variants and reduction may be in free variation. A good example of this sort comes from the realizations of the preposition *em* ‘in’. As shown below, there are cases where only allomorphy, but not reduction, may account for certain alternations. This is the case of the alternation [ẽj] / [n]. In general, only one of these two forms is possible, as in (482a) and (482b). However, both lexicalized and non-lexicalized forms freely coexist in particular contexts, as illustrated in (482c). Observe, additionally, that whenever [ẽj] is possible, the reduced variant [ẽ] is also allowed, together with intermediate realizations.

- | | | |
|--------------------------|----------------------------------|----------------|
| (482) a. non-lexicalized | em Agosto | [ẽj],[ẽ] |
| | ‘in August’ | |
| b. lexicalized | na | [n] |
| | in-the-FEM | |
| c. free variation | em algum / nal gum | [ẽj],[ẽ] / [n] |
| | in some-MASC | |

As mentioned in section 5.4.2 above, the reduced lexicalized form [n] is probably a remnant of an earlier stage of the language when the preposition’s vowel could be deleted, possibly as a consequence of reduction, and

the underlying consonant surfaced as the onset of the following onsetless syllable. Since this type of resyllabification with nasal segments is no longer possible at the present stage of Portuguese, it is clear that only lexically stored allomorphs may account for these alternations. The reduced variant [ẽ], by contrast, may be viewed to result from reduction. It is important to remark, nonetheless, that we cannot determine at this point whether or not [ẽ] has already lexicalized as an allomorph of *em* that freely coexists with [ẽ̃].

The difficulty in determining the lexical status of reduced variants may further be illustrated with consonant-schwa clitics, as shown in the following discussion on the preposition *de*. As we have already seen, when this preposition is followed by a word that starts with a vowel, the clitic's vowel may in general either semivocalize or disappear. In special contexts, by contrast, only the form without the glide is allowed. In the latter cases, it is clear that lexicalization has occurred (see section 5.4.2). In the former, by contrast, we may either admit that /de/ and /d/ are two allomorphs of this preposition, or hypothesize that vowel deletion is simply the extreme of a reduction continuum.³⁷⁸ Pointing to the allomorphy solution is the fact that, in the sequence glide-vowel, speakers only judge possible glide-deletion if the glide belongs to a clitic, but not, for instance, if both the glide and the vowel occur word internally. However, a more careful examination of EP speakers productions reveals that very reduced almost absent glides do occur at least within certain words, such as *reunião* [Rjunjãw]/[R^hjunjãw] 'meeting'. The fact that, unlike what happens with schwa-clitics, vowel deletion in these cases is judged to be impossible may suggest that here we are facing reduction, while in schwa-clitics we are in the presence of allomorphy (together with reduction in the intermediate cases). However, the high frequency of consonant-schwa clitics, together with their higher rate of reduction may also explain speakers' different judgements. Above all, what these observations show is that it is crucial to know more about phonetic reduction in order to understand what is specific of clitic words and what is found more in general in the language (although, possibly, less frequently). In particular, it seems essential to gather knowledge based on spontaneous speech and more informal registers, and to determine with accuracy the limits and probability rates of reduction, as well as the factors that (dis)favor it. Such an enterprise, however, goes well beyond the purpose of the present book.

Independent evidence that allomorphy and reduction must be allowed to coexist, is provided by the observation that items may show too many in-

intermediate forms for us to admit that each of them is stored in the lexicon. This is particularly well exemplified by the English realizations of the article *the*. It is generally accepted that this item has two lexically stored allomorphs, namely [ðə] and [ði]. Nevertheless, as pointed out in Jurafsky, Bell, and Girand (2002), a corpus like the Switchboard (Greenberg, Ellis, and Hollenback 1996) contains 33 different phonetic transcriptions of this article. These surface realizations appear to be dependent on the segmental and prosodic context, as well as on other factors such as speaking rate. It is highly unlikely that all such realizations correspond to lexically stored allomorphs.³⁷⁹

We have seen that lexically stored allomorphs may be in free variation. In such cases, we suggest that factors like the position within the intonational phrase, or register, may favor the selection of (un)reduced forms. This is, in fact, entirely compatible with the observation that weaker variants can also be found in I-initial position, and that some may even bear I-initial prominence.³⁸⁰ Indeed, if all reduced variants were merely the result of phonetic reduction, all such elements should be excluded from actually bearing prominence. Since the clitic variants that bear stress should not to be affected by phonetic reduction, prominence assignment may constitute a useful diagnostic for the identification of lexically stored allomorphs.³⁸¹

To sum up, we have seen that both reduction, understood as a phonetic gradient phenomenon, and allomorphy may be involved in the realizations of the clitic words studied in this chapter. However, the ambitious task of determining which clitic forms are lexically stored and which result from reduction is beyond the limits of the present study. For the purposes of this book, it was important to distinguish the phenomena that affect clitic words from the general phonological processes that are found in EP. The observation that clitics may undergo specific kinds of reduction led us to put forth the hypothesis that some clitic forms are lexically stored as allomorphs. The fact that clitics show gradient realizations was seen, in turn, to result from phonetic reduction, which may also affect lexically stored allomorphs. Contrasting allomorphy with reduction, we have suggested that (i) speakers tend to be aware of the shape of lexically stored allomorphs; (ii) *weak* allomorphs may bear (postlexically assigned) prominence; (iii) speakers tend not to be aware of non-lexicalized variants; (iv) reduced variants are gradient; (v) reduced variants may not bear prominence; and (vi) very high frequency may favor speakers consciousness of reduced variants, as well as their lexicalization. Finally, allowing both allomorphy and reduction was seen to be compatible with the observation that factors such as I-initial position disfavor the occurrence of reduced forms: in the cases involving al-

lomorphy, I-initial position may be seen to favor the selection of stronger allomorphs; in the cases resulting from reduction, given that I-initial position is often assigned prominence, reduction is expected to be often inhibited in this position.

The issues raised in this section concerning phonetic reduction and the separation of reduction from allomorphy are certainly fascinating. Many already foreseen ramifications have been deliberately left untouched. We believe, in fact, that the complexity of this area of research deserves a much thorough investigation, which will have to be left for future work.

6. Summary of main findings

The primary goal of this chapter was to demonstrate that the phonological shape of clitics cannot always be derived from the application of regular processes of EP. In section 5.1, a systematic comparison between sequences involving clitics and parallel sequences involving prosodic words has showed that the phenomena under investigation in this chapter are not general in the language. It was seen, nevertheless, that specific items that have the properties of prosodic words are also subject to the same type of reduction phenomena. The common feature these words share with clitics is that the elements of both groups belong the set of high frequency words of EP. This led us to conclude that the reduction phenomena observed in clitic words are not specific to unstressed items but must result from a frequency effect instead (section 5.2).

With respect to the shape of clitic words, the occurrence of intermediate forms was seen to point to the application of gradient reduction, since the existence of lexically stored allomorphs would seem to imply categorical realizations (section 5.1). Nevertheless, the discussion in section 5.5 has revealed the benefits of admitting that some clitic forms may result from lexically stored allomorphs. Like other elements of the language, allomorphs may also be subject to phonetic reduction, thus explaining some of the gradient effects found.

Considering the reduced forms of clitics and other high frequency words, we were able to establish the following main generalizations: (i) the items that clearly show reduced forms constitute a subset of the set of EP highly frequent words; (ii) reduction usually affects phonological material of word initial syllables; (iii) most reduced forms exhibit syllable structures that are less marked than those of unreduced forms; (iv) reduction of disyllabic clitics usually results in monosyllabic forms (section 5.3).

In order to investigate the factors that condition the realization of (un)reduced clitic forms, we have collected data produced by three speakers of EP (section 3). The results have shown that, for each function word studied, I-initial position clearly disfavors the occurrence of reduced forms, when compared to the I-internal position. Other factors found to constrain the occurrence of (un)reduced forms are: (i) speech rate, as reduced forms occur more often in faster rate when compared to normal rate; (ii) speakers, since some speakers tend to use reduced forms more often than others; and (iii) word category, for complementizers tend to reduce less often than prepositions and pronouns (section 4).

The effect of I-initial position on the shape of clitic words was seen to indicate that this is a prominent position in EP (section 5.4.1). This was supported by independent data, showing that perceived non-primary stresses in EP are noticeably more frequent in I-initial position than in I-internal position. The prominence features associated with I-initial position were further conjectured to be responsible for the generalized appearance of postverbal pronominal clitic location in EP, which contrasts with other Modern Romance languages in this regard. It was hypothesized that I-initial prominence has emerged at the time when pronominal clitics became preferably postverbal, and that this clitic location has constituted a strategy that allowed clitics to escape from occurring in such prominent positions.

Finally, word category was found to be relevant both for the rate of reduction and for the processes of lexicalization of adjacent function words (section 5.4.2). It was shown that the same elements that exhibit a higher rate of reduction, prepositions in particular, are those that participate in lexicalized expressions with items that belong to closed classes, and with which they appear very frequently combined. The fact that complementizers reduce less often than prepositions, in turn, was proposed to follow from an inherent property of complementizers, which is likely to result from the fact that these items usually occur in prominent positions, namely, at the edges of the intonational phrase domain.

As a general conclusion, we have shown that the phonetic shape of EP clitic elements is often not explained by the application of general phonological rules of the language. We have suggested that phonetic reduction, together with allomorphy, must be crucially involved in the account of the variants of clitic words. However, given the number of factors that may constrain reduction and the fact that a thorough account of phonetic reduction in EP is clearly out of the scope of this book, we have left open many of the interesting questions that may be raised in this domain. This is therefore a topic to develop in future work.

³²⁹ In our terms, unaccented words are lexically unstressed items, which may undergo reduction processes already within the lexical component, and whose reduced forms may thus not be entirely explained by their occurrence in unstressed position postlexically, unlike simple clitics.

³³⁰ In Hall's (1999a) data, full vowels are always reduced to schwa in the weak forms, except if they are deleted, which usually happens when they are followed by another vowel. In addition to this rule, German articles are also argued to be subject to a rule that deletes initial consonants (cf. Wiese 1996: 7.4.3).

³³¹ After controlling for a number of variables, Jurafsky, Bell, and Girand (2002) conclude that the differences noticed result from word-predictability factors, which include relative frequency.

³³² Notice that, according to Keating this difference disappears when low-frequency content words are considered.

³³³ On the basis of these results, it is suggested that presence/absence of reduction may be used by speakers to signal new/old information (Bybee 2000: 253).

³³⁴ We should point out that the facts just described distinguish vowel reduction in EP from vowel reduction in other languages. For example, according to Van Oostendorp (2000: 131), in English if a vowel can reduce, reduction is always optional. In Dutch, reduction is also usually optional (Booij 1995; Van Oostendorp 2000). Notice that some Dutch vowels always surface reduced, and thus are assumed to be underlying schwas. In other cases, unreduced and reduced vowels alternate in morphologically related forms, but reduced forms are obligatory. These cases are assumed by Booij (1995) to result from allomorphy.

³³⁵ In an acoustic study on segments deletion in EP, Delgado Matins (1975) reports that [v] is deleted in 7 out of a total of 442 vowels. Unfortunately, no information is given with respect to the total number of occurrences of this particular vowel or the context of these deletions.

³³⁶ In Brazilian Portuguese, where vowel deletion as a consequence of reduction processes does not occur, there is also vowel deletion due to the resolution of across-words hiatus (e.g. Bisol 1992, 1996; Abaurre 1996).

³³⁷ Again, in Brazilian Portuguese [i] in unstressed positions is also subject to deletion processes conditioned by the segmental context. For example, this vowel like other non-low vowels may be deleted between homorganic consonants (cf. Bisol and Hora 1994).

³³⁸ The phonetic transcription of *devedor* and *despreparar* in (5a) are taken from Mateus and d'Andrade (2000: 44); the transcriptions of *professor* and *docente* are based on acoustic data presented in Vigário (1998a: Appendix II) and Mateus and Delgado Martins (1982), respectively.

³³⁹ According to E. d'Andrade, a personal communication cited in Andrade (1980: 67), [i] truncation is less favored after a word initial consonant (which is the only word initial context where schwa may occur). This may be seen to follow from the

prominence properties of the prosodic word initial position. Recall that this position may be assigned initial and emphatic stress, it defines the context for the application of initial /r/-strengthening, and vowels in prosodic word initial position do not undergo full (lexical) vowel reduction (see in particular section 3.9).

³⁴⁰ Notice, however, that this analysis was abandoned in more recent work (e.g. Mateus and d'Andrade 2000).

³⁴¹ Repetitions were allowed and all natural-sounding renditions were considered for analysis.

³⁴² Complementizers in intonational phrase final position, which surface with a schwa, were eliminated from this analysis. A few renditions read with hesitations were also not considered.

³⁴³ The importance of vowel quality and stress in the following vowel for schwa deletion is suggested in Gonçalves Viana (1883: 39–40). However, our observations did not confirm the relevance of these variables.

³⁴⁴ For this preliminary observation we have considered the data of two of the three speakers whose data is analyzed in the present chapter. The data of another speaker was also analyzed. Because this subject had many hesitations in the reading task and she became unavailable for the recording of the second set of materials we have decided to eliminate her renditions. We should point out, nevertheless, that the eliminated data showed precisely the same tendencies as the data described below (see Vigário 1998b).

³⁴⁵ In a study conducted simultaneously, Frota (2000: 5.2.3.1) reports the same tendency for the function word *aos* (to-the-MASC-PL) ([awʃ]/[ɔʃ]) to reduce less in intonational phrase initial position when compared to I-internal position. On the reduction of this function word, see also our results in section 7.3.2 below.

³⁴⁶ No noticeable differences were found in the renditions of the different texts. This does not mean that register is not relevant for schwa deletion or other reduction phenomena. Rather, we believe that, in general, speakers have probably employed a formal register during the reading sessions. This is supported by the inspection of the productions of other function words in the task of retelling the tale. Some reduced forms that intuitively are style dependent, like [pə] (<*para*) and [ku] (<*com o*), occurred predominantly in this task.

³⁴⁷ We have eliminated from the data analyzed a few cases where the position of clitic function words could not be undoubtedly ascertained due to unclear cases of intonational phrase boundaries.

³⁴⁸ The reason for this may be stylistic. For some speakers the unreduced form of *pelo(s)/pela(s)* is felt to belong to rather marked registers. The existence of the nominal form *pelo* [pɛlu] 'hair' may also contribute for the preference for the reduced form of the function word.

³⁴⁹ We should point out that the deletion of the vowel of the preposition, which may also occur with the article *a*, is usually felt to belong to very marked registers. As these examples show, however, it also occurs in other types of speech, though

possibly less frequently. We should add that in 3 tokens realized by the same speaker a denasalized form of *com* was found followed by a full word. The number of cases of this sort, however, seems negligible and does not undermine the generalization that reduced forms of *com* are typically found when followed by articles.

³⁵⁰ In this respect, speakers seem to behave differently. IS has never realized the relevant vowel as a glide when the following vowel is stressed. We also tend to follow this pattern. We should point out that, although the presence of the glide in this context is reminiscent to the presence of the glide in stressed environments within compound prosodic words (cf. chapter 6), the two cases are not identical: while the glide is obligatory when the following vowel bears the compound prosodic word prominence, in this case, the absence of the glide is possible, occurring in 57% of the realizations of RI and IV where the relevant context obtains.

³⁵¹ Recall that the list of Bacelar, Marques, and Segura da Cruz (1987) includes the most frequent words of EP, based on a total of ca. 700 000 collected words.

³⁵² The word initial fricative may also be deleted in the inflected forms of the verb *estar* ‘to be’, possibly with the exception of the forms of the future and conditional, where deletion seems a bit more marked.

³⁵³ In the more reduced form we believe only the meaning ‘great’ is possible. We should add that this form is also subject to a process that consists of the change of a schwa-vowel (or sometimes [u]) into [ɐ], which is usually associated with very marked/young people’s registers and typically occurs in imperative or exclamative sentences, as in the examples below:

Que grande parvo!	[kigrẽdparvu]/[kɛgẽdɐparvu]	‘What a great jerk!’
Vou-te contar!...	[vótkõtár]/[vótɛkõtár]	‘I’m telling you!...’
Ele é muito giro.	[élémũjtuzĩru]/[élémũjtɐzĩru]	‘He is very cute.’
Jesus!	[ʒizũ]/[ʒɐzũ]	‘Jesus Christ!’

³⁵⁴ Word final fricative deletion also occurs with the clitic *mas* [mɐʃ] ‘but’ when followed by *não* ‘not’.

³⁵⁵ Our contention is not that reduction may not affect these words, but rather that these forms are judged to be ill-formed, contrary to those in (28) and (29a–31a). Whether the contrast noticed follows from allomorphy or from the fact that reduction may simply be more frequent in the words in (28) is a question we leave open (see also the discussion in section 7.4.5). Notice, additionally, that the observed differences could be thought to be related to the content versus functional nature of the relevant words. However, most items in (34) are function elements. Furthermore, Jurafsky et al. (2001) show for English that content words may also be subject to different amounts of reduction, depending on their relative frequency.

³⁵⁶ The negative word *não* [ˈnẽw] ‘not’ may also show different realization, in this case depending on the segmental context. For example, the negative sentences

presented in Vigário (1998a) (see in particular Appendix II), include sequences such as *não emprestaram* ‘(they) did not lend’ and *não ofereceram* ‘(they) did not offer’, which are in general transcribed as [nẽwẽpʃtárẽw] and [nẽwõfsérẽw]; however, the transcription [nẽçpʃtárẽw] and [nõofsérẽw] is also found (e.g. sentences 33 and 32 of Appendix II, respectively) (in these transcriptions, primary word stress was only marked in words with more than one syllable). Notice that in all such cases, the negative word was realized with a stress that makes the whole word to become perceptively highlighted. Thus, here we have evidence that the emergence of these forms is not a consequence of destressing. The prosodic status of negative words, and in particular of the sentence negation particle, requires further investigation, which will be left for future work.

³⁵⁷ Diachronic data provide additional examples of possible reductions occurring in similar environments. For example, in the evolution of the form of Latin pronouns that originated the definite articles and accusative pronouns *o(s)/a(s)* the-MASC(PL)/FEM(PL) of Portuguese, the pronoun’s initial syllable has been deleted in the whole paradigm. This is shown in (i), taken from Williams (1975).

(i) Classic Latin	Vulgar Latin (reconstructed)	
illum (accusative, singular)	lo	(masculine, singular in Portuguese)
illam (accusative, singular)	la	(feminine, singular in Portuguese)
illos (accusative, plural)	los	(masculine, plural in Portuguese)
illas (accusative, plural)	las	(feminine, plural in Portuguese)

The items *no(s)/na(s)* in-the-MASC(PL)/FEM(PL), formerly, *eno(s)/ena(s)* (cf. Williams 1975: §137.4) illustrate the same point.

³⁵⁸ In her discussion on the clitic status of the Italian pronoun *loro*, Nespor (1999a: 875) explicitly raises and leaves open the question of whether disyllabicity may be used as a diagnostic against the phonological clitic status of a word. That is, clitics are conjectured to be at most monosyllabic universally, as suggested by Neijt (1985) and Berendsen (1986: chap. 2, note 2), among others. In EP, there is evidence that at least four disyllabic words are clitic – *para* ‘for’, *pelo(s)/pela(s)* by-the-MASC(PL)/FEM(PL), *porque* ‘because’ and *cada* ‘each’. Nonetheless, EP follows the general tendency in that the great majority of clitic words are at most monosyllabic. We should add, furthermore, that the reduction of disyllabic clitics via the deletion of the vowel of the first syllable results in the creation of a complex onset. Interestingly, in the case of *para* the complex onset eventually becomes simple through consonant deletion.

³⁵⁹ In accordance with our analysis of proclitic function words (see chapter 5), we are assuming that stress in the proclitic counts as a prosodic word initial stress. Since such stress may only occur once per prosodic word, whether or not with prosodic adjuncts, the number of potential prosodic word initial positions relevant for initial stress is identical to the number of prosodic words in the corpus.

Prosodic words starting with a syllable bearing word stress and not preceded by a proclitic are, naturally, excluded.

³⁶⁰ According to Martins (1994: chaps.3–4), postposition of clitics is the dominant pattern in the 13th century, whereas preverbal clitic location becomes almost generalized in the 15th century. Nevertheless, free variation in clitic positioning is still attested in 16th century.

³⁶¹ The same general observation is made in Galves and Galves (1995) and Britto (1998), although the shift is considered to be more recent. The precise point in time where this change took place is not relevant for our discussion.

³⁶² On the phonology of enclisis and proclisis in Old French see also Jacobs (1993).

³⁶³ The relevance of prosodic information in the distribution of clitics is also proposed in Halpern (1995), Taylor (1996), and Pintzuk (1996) for languages such as Serbo-Croatian, Bulgarian, Old French, Ancient Greek, Old English, among others. According to Booij (1996a), the Dutch pronoun *ie*, unlike other clitics in the language, also requires a prosodic host to its left, and thus may not occur sentence initially (see also section 1.5.3 of this book).

³⁶⁴ Notice furthermore that three sentences of the BP corpus were excluded (see the preceding note), and thus the PB corpus is slightly shorter than the EP one. According to Frota and Vigário (2000), the perceived prominences in PB follow an alternating strong-weak pattern. Additionally, unlike in EP, in BP prosodic word internal non-primary stresses are very frequent, occurring in 69.7% of the cases where the context for the alternating pattern obtains.

³⁶⁵ Three sentences read with hesitations were excluded from the materials.

³⁶⁶ For the purposes of this argumentation we are assuming, with Galves and Galves (1995), that the relevant prosodic properties of BP are similar to those of Classical Portuguese spoken before the change took place in the European variety. Other prosodic and rhythmic differences between the two varieties of Portuguese, some of which bearing on the phonological properties of the intonational phrase, are described in Frota and Vigário (2000, 2001).

³⁶⁷ There are a few cases where both the lexicalized fused form and the regular non-lexicalized sequence may occur, as illustrated below (see also section 7.4.5):

em um, em uma	num, numa	in a-MASC, in a-FEM
em outro, em outra	noutro, noutra	in other-MASC, in other-FEM
em algum, em alguma	nalgum, nalguma	in some-MASC, in some-FEM

³⁶⁸ *pelo(s)/pela(s)* have lexicalized in a period when the form of the preposition that became *por* was *per* and the form of the article after a word ending in consonant was *lo(s)/la(s)*. The deletion of the preposition final consonant is a process similar to the one that gave rise to the precompiled rule involving verbs ending with consonant when followed by the accusative pronoun with the form *lo(s)/la(s)* (see section 4.3.2.2). On the history of these forms, see Williams (1975: §143).

³⁶⁹ Notice, nonetheless, that sequences of pronominal clitics may also lexicalize, as we have seen in section 4.3.2.2. We should recall that pronominal clitics behave like prepositions in showing a rate of reduction higher than complementizers.

³⁷⁰ Fusion with *com* has in fact occurred in personal pronouns, as shown in (i) below. This case is nevertheless slightly different, since the forms *migo*, *tigo*, *nosco*, *vosco*, which were formed in Latin (*me*, *te*, *nos*, *vos* plus *cum*) had lost their original meaning and were not used as independent words in other contexts when fusion took place, unlike in all the cases reported above.

- | | | | |
|-------------------|----------------|-----------------|----------------|
| (i) <i>comigo</i> | 'with me' | <i>connosco</i> | 'with us' |
| <i>contigo</i> | 'with you(SG)' | <i>convosco</i> | 'with you(PL)' |

³⁷¹ In our corpus, we only have 4 cases of *sem*. In all the realizations of these cases the three speakers, this word was pronounced as [sẽj]. We believe, in fact, that the reduction of this function word is totally impossible. We should add that, although *em* is much more frequent than *sem*, according to the EP frequency list in Bacelar, Marques, and Segura da Cruz (1987), *sem* is among the 140 most frequent words of EP. It is conceivable that the reason why this word does not undergo reduction is because it bears word stress, unlike *em*, but we do not have independent evidence in favour of this hypothesis at the present moment.

³⁷² In fact, the same conditions for lexicalization may also obtain with *que* in expressions such as *se bem que* 'although' (lit. 'if well that'), *ainda que* 'even though' (lit. 'still that'), *já que* 'since' (lit. 'already that'), *só que* 'although' (lit. 'only that'). In a few cases, when such expressions occur in intonational phrase final position, *que* may surface without stress, similar to what obligatorily happens with pronominal enclitics (see section 5.3.3.1). Since this behavior is exceptional, it may be seen to follow from lexical information. Notice, nevertheless, that the exceptional form involving the enclisis of *que* must coexist with the regular form, since this complementizer may also behave like other proclitic words (see also section 7.4.5 for the evidence that lexically stored allomorphs may be in free variation with non-lexicalized units).

³⁷³ This difference is noticed by EP speakers, which often judge the reduced forms of consonant-schwa complementizers to denote a very marked register, unlike in the case of reduced prepositions. The different behavior of the two categories may already be inferred from the description of Gonçalves Viana (1883).

³⁷⁴ In order to have comparable data, we have excluded from the calculation of Bacelar, Marques, and Segura da Cruz data the cases where *de* has lexicalized with the following function word.

³⁷⁵ We should point out, additionally, that *de* may also function as a complementizer and thus some of the cases listed in Bacelar, Marques, and Segura da Cruz (1987) under this form are probably complementizers. The actual number of occurrences of *de* functioning as a preposition is therefore likely to be slightly lower. Notice further that *que* may also function as an interrogative word (both in

the frequency corpus of EP, and in the corpus we have analyzed). Its occurrence as a complementizer is, nevertheless, much more common in the language.

³⁷⁶ In Frota and Vigário (1996) and Vigário and Frota (1998) a hypothesis is put forward that the set of items that function as proclisis triggers in EP is the set of strong function words. The facts reported in this chapter, and in previous work (Vigário 1999a), add to this idea, since they show that complementizers behave like elements stronger than prepositions, in that they reduce less and often appear in I-prominent positions. Recall that complementizers, but not prepositions, are proclisis triggers in EP.

³⁷⁷ It should be noticed, nevertheless, that precompiled allomorphy can also be hypothesized to be dependent on phrasal prosodic information (e.g. Nespor 1990) and on speech register.

³⁷⁸ Notice that the realization of the glide is a consequence of the application of a general process that affects the first unstressed vowel of a VV sequence, which is V1-semivocalization (cf. section 3.10.2).

³⁷⁹ The same observation applies to other English function words, as Jurafsky, Bell, and Girand (2002) list 82 common pronunciations of the 10 most frequent function words of English (including 9 variants of the article *the*).

³⁸⁰ For example, the reduced form [prɐ] of the clitic *para* 'to' is attested in our corpus in I-final stressed position. I-initial stress in this form was also occasionally signaled.

³⁸¹ Unfortunately, we did not systematically investigate the assignment of prosodic word or intonational phrase initial prominence in our corpus, nor did we transcribe the stressed forms that occurred in I-final position. Therefore, we cannot determine with precision which clitic forms may receive such stresses. The identification of the clitic forms that may bear stress, which we propose should not follow from phonetic reduction, is thus left for future research.

Chapter 8

Conclusion and directions of future research

1. Introduction

This chapter contains the main conclusions reached in this book. Some of the implications of the analyses proposed are also highlighted and a few additional questions are raised pointing to directions of future research.

One of the main goals of this work was to isolate the phonological properties that define the prosodic word in European Portuguese. We have seen that evidence for the prosodic word in this language is abundant, as we were able to determine a number of phonological phenomena that refer to the prosodic word and consequently constitute reliable diagnostics for this prosodic domain in EP. Not taking into consideration the distinction established in chapter 3 between lexical and postlexical phenomena, these include (i) several phonological processes, (ii) prominence phenomena (iii) tonal phenomena, (iv) phonotactic restrictions, and (v) constituent deletion processes (see below for a more detailed list). By contrast, other types of diagnostics, such as minimal word size requirements, argued to play an active role in many languages (see chapter 1, section 6.1), were shown in chapter 5, section 2 not to operate in EP.

2. Minimal prosodic words

Given the existence of sound cues not only for determining the presence of prosodic words, but also for identifying both edges of this prosodic domain, and always on the basis of such evidence, we were able to establish the morphosyntactic units that are grouped together within a minimal prosodic word in EP, that is, a prosodic word that immediately dominates the next lower constituent in the prosodic tree (chapter 5). We have seen that a minimal prosodic word includes either a stem plus suffixes that do not constitute independent domains for word stress assignment, or affixes (prefixes or suffixes) that constitute independent domains for word-level stress. We concluded, furthermore, that in the same way suffixes that do not form an independent stress domain are incorporated into the same prosodic word as their morphological base, pronominal enclitics are incorporated into the

prosodic word that dominates their host. As for unstressed (non-lexicalized) prefixes, as well as proclitics, they were shown to behave phonologically like prosodic word initial elements. Since the initial segments of their base/host also exhibit the phonology of prosodic word initial units, and adopting the view on the prosodization of clitics expressed in Selkirk (1996), Booij (1996a), and Peperkamp (1997a), among others, we have proposed that prefixes and proclitics are adjoined to (rather than incorporated into) the prosodic word that dominates their base/host, yielding a recursive structure.

In the light of these findings, we have proposed in chapter 5, section 6 and chapter 6, section 5.3 that the lexical prosodic word formation in EP is obtained via (a) a morphology-phonology mapping condition, stating that a stem plus following suffixes form a prosodic word (like, for example, Latin, Greek, Italian, Spanish, Hungarian, and French – cf. Nespor and Vogel 1986; Vogel 1990; Peperkamp 1997a; and Hannahs 1995b); (b) a general well-formedness condition limiting the number of word stresses per minimal prosodic word to one, which is responsible for the prosodic word status of affixes that form independent stress domains; and (c) a parsing condition stating that morphemes that are not dominated by the prosodic word are adjoined to the following prosodic word, which is responsible for the prosodization of (unattached) stressless prefixes.³⁸²

3. Clitics and their postlexical attachment

EP emerges as a language with a clear tendency for proclisis – except for postverbal pronominal clitics, all other clitic elements display the properties expected from proclitic objects. The fact that only postverbal pronouns are enclitic in this language could lead us to cast doubt on the clitic status of these elements. The hypothesis could be entertained that clitic pronouns have been reanalyzed as lexicalized inflectional affixes (as proposed, for instance, in Zwicky 1987a and Halpern 1995). Such a view would explain why these elements, and only these, are incorporated into the verb's prosodic word, thus behaving in a way similar to regular suffixes. However, this hypothesis was rejected on the basis of a number of arguments presented in chapter 4. Although a few facts superficially suggest the lexical attachment of pronominal clitics, the bulk of the evidence gathered showed that clitic pronouns must attach to their hosts postlexically. Such evidence includes the position of clitic pronouns with respect to the verb, their scope in coordinate structures, the possibility of limited interpolation, and the be-

havior of clitic pronouns with respect to many phonological processes. As for the evidence for lexical attachment – namely, mesoclitisis, idiosyncratic phonology, and host selectivity – it was claimed to be only apparent. With respect to mesoclititic structures, we believe to have brought to light important arguments against the lexical status of this construction. As for the idiosyncratic phonology that verb-host sequences display, we have proposed that it follows from lexically stored information (precompiled rules and allomorphy), along the lines of precompilation theory (Hayes 1990; Nespor 1990). Finally, the fact that clitics attach to X^0 -level units was seen to follow naturally from a theory of cliticization (or phrasal affixation) as the one proposed by Andrews (1992).

Under our proposal, the exceptional behavior of postverbal enclitics with respect to the direction of cliticization follows from the interaction of two conditions: the phonological dependence of these stressless units; and the syntactic configuration where they occur, ending under the same syntactic head as the preceding verb (cf. Duarte and Matos 2000). In fact, as far as we could determine, in no other circumstance do these two conditions coexist, and therefore we do not expect to find other instances of phonological encliticization in EP. In contrast, we may observe what happens when each of these conditions occurs independently. Let us take the syntactic configuration first. When two stressed units are found together under a single syntactic head, as undoubtedly is the case of morphological compounds, phonological evidence indicates that they are mapped onto a single (compound) prosodic word. This clearly suggests that what is crucial for the prosodization within a single prosodic word of either sequences of verb+clitic pronouns or units like morphological compounds is the syntactic configuration they display. The difference between these units, on the other hand, follows exclusively from their phonological properties: clitics lack stress, and therefore become phonologically dependent on the prosodic host, whereas (non-lexicalized) morphological compounds are composed of two or more prosodic words, thus giving rise to a compound prosodic configuration. The second condition, that is the occurrence of phonological dependence without the relevant syntactic configuration, is illustrated by proclitic function words, such as articles or prepositions. These elements are stressless units that do not occupy the same syntactic head as the preceding lexical unit (if any). Similar to postverbal stressless pronouns, proclitics are prosodized in such a way that they become dependent on an adjacent prosodic word. The difference between proclitics and enclitics resides in the type of prosodic relation of dependence that each of them establishes with their phonological host. In the case of proclitics, phonologi-

cal evidence shows that they adjoin to the prosodic word that dominates the phonological host, rather than incorporate into it, like enclitics do.

In our view, the asymmetric prosodization of enclitic and proclitic function words in EP is explained by the interplay between lexical and postlexical information. Postlexically, only the left-edge of lexical prosodic words appears to be preserved, while the right-edge becomes postlexically invisible, when it does not coincide with the right-edge of a syntactic terminal node. In order to capture these observations, we have proposed that only the left-edge of lexical prosodic words is projected in the postlexical construction of prosodic words (see chapter 5, section 6 and chapter 6, section 5.3). The same mechanism was also shown to account for the behavior of prefixes and suffixes that do not form independent stress domains, which exhibit a postlexical phonology similar to proclitic function words and enclitic pronouns, respectively.

An immediate consequence of our findings is the rejection of previous proposals defending that EP tends towards leftward cliticization (e.g. Brandão de Carvalho 1989; Galves and Galves 1995). The question could then be raised of why some languages tend to encliticization and allow bidirectional cliticization, like Dutch and German (cf. Booij 1996a and Kleinhenz 1996, respectively), while others, like EP, tend to procliticization and do not allow for bidirectionality. We have attributed the EP requirement for rightward attachment to a general condition imposed on this particular language (see chapter 5, section 6 and chapter 6, section 5.3). Nevertheless, it would be desirable to derive such condition from more general properties of the language, such as syntactic or rhythmic information. As the brief discussion in the following paragraph shows, there is not an easy answer to this question.

Other Romance languages, such as Italian and Spanish (cf. Peperkamp 1997a), or Brazilian Portuguese (cf. Bisol 2000; Schwindt 2000), similar to EP, are not reported to admit bidirectionality. Except for cases involving pronominal clitics, like in EP, enclisis is not referred to be allowed in these languages. From a rhythmic point of view, EP displays some properties noticeably distinct from other Romance languages (cf. Ramus, Nespor, and Mehler 1999 for a rhythmic classification of Italian, French and Spanish and Frota and Vigário 2000 for rhythmic differences between European and Brazilian Portuguese). These rhythmic differences appear, nevertheless, not to impinge on the direction of cliticization, contrary to what is suggested in Galves and Galves (1995). In contrast, EP is syntactically much closer to Italian, Spanish, and Brazilian Portuguese (BP), than it is to Dutch and German. Furthermore, the latter two languages are typologically akin, dis-

playing, for example, a heterogeneous system in the head-complement order, in addition to structures with function elements closing syntactic phrases. Since these properties do not characterize Romance languages, this could be taken to suggest that the difference between the two groups of languages concerning clitic directionality may result from syntactic information. However, a parallel contrast obtains when we consider other phonological properties, such as stress location at the level of the word and higher domains. While in Romance languages, EP included, stress is rightmost in all prosodic constituents from the level of the prosodic word upwards, in German and Dutch words may be initially stressed (although word stress is in general computed from the right-edge of prosodic words), and compound structures display both weak-strong and strong-weak patterns, whereas ϕ s are in general righthanded (e.g. Booij 1995; Wiese 1996; Visch 1999; Zonneveld et al. 1999; Nespor 1999b; and chapter 6, section 5.2). An additional source of contrast between the two groups of languages concerns the importance of phonotactic constraints in the formation of prosodic words. While in Dutch and German phonotactic constraints may be responsible for the prosodization of suffixes as prosodic words (cf. Booij 1996a; Wiese 1996), and in some cases are argued to have implications in the prosodization of clitics (cf. Hall 1999a for German), this appears never to be the case in languages like Italian, Spanish, BP and EP (e.g. Nespor and Vogel 1986; Peperkamp 1997a; Bisol 2000; and chapter 5 of this book).³⁸³ Which of these factors, if any at all, is responsible for language variation with respect to the direction of cliticization is an open issue.

4. Compound prosodic words

Phonological evidence has also led us to propose in chapter 6 that a number of morphosyntactic units are grouped together under a prosodic node that includes two or more prosodic words, but that does not display the properties of a phonological phrase. The behavior of such a domain has compelled us to analyze it as a compound prosodic word, that is, a domain that is immediately dominated by the phonological phrase and that dominates two or more prosodic words. The units which form compound prosodic words in EP comprise (i) derived words with suffixes that constitute stress domains independent of their base, (ii) derived words with stressed prefixes, (iii) morphological (stem plus stem) compounds, (iv) morphosyntactic (V+N) and lexicalized syntactic (word plus word) compounds, (v) abbreviations,

(vi) mesoclititic structures, and (vii) sequences of prosodic words consisting of (a) pairs of letter names, (b) names of letters followed by numerals, as well as (c) (some) numerals followed by the words *horas* and *anos*. Inspired by other proposals, such as Nespor and Ralli (1996), Peperkamp (1997a), and Nespor (1999b), we have interpreted the prosodization of such constructions to follow from syntactic information, since all of these constructions appear to occupy a single syntactic terminal node. In particular, we have accounted for their prosodization via a *syntax-phonology mapping* condition according to which X^0 -level expressions are mapped onto a prosodic word in EP (see chapter 6, section 5.3).

All the complex words that were previously found to have internal prosodic words in a language like Italian (cf. Peperkamp 1997a) were found to form compound prosodic words in EP, namely, derived words with stressed prefixes, morphological (/stem+word) compounds, and V+N compounds. Given that in the two languages such units display the same grammatical properties that are relevant for the construction of compound prosodic words, we suggest that our analysis may be extended to Italian. We have also discovered other types of structures that display the phonology of compound prosodic words – specifically, derived words with suffixes that constitute stress domains independent of their base, “mesoclititic” structures, abbreviations, and a limited number of expressions involving names of letters and/or numerals. Like the constructions previously referred, all these structures occur under a syntactic terminal node that includes units that form more than one prosodic word. With respect to mesoclititic structures and derived words with suffixes that form stress domains independent of their base, they do not exist in Italian. As for abbreviations, we may speculate that they are likely to form compound prosodic domains in Italian as well, given the similarity between the two languages with respect to prosodic word formation.

The fact that the kind of morphosyntactic structures that are prosodized as compound prosodic words in EP coincide at least partially with the units that display a strong-weak stress pattern in Germanic languages such as English and Dutch, has led us to suggest that the compound prosodic word analysis could also be extended to these languages (chapter 6, section 5.2). In contrast to the weak-strong pattern that characterizes the phonological phrase in these languages, a number of word compounds display a strong-weak pattern (e.g. Booij 1995; Nespor 1999a; Visch 1999). Furthermore, in Dutch the compound strong-weak stress is found also in derived words comprising suffixes that form their own prosodic words (Booij 1995). It was shown that EP provides clear phonological evidence that constructions

of the latter type form compound prosodic words (chapter 6, section 3.1). Under our proposal, the data from Dutch and English could result from the fact that these units, like the relevant elements in EP (and possibly in Italian), occupy a single syntactic terminal node, which is mapped onto a prosodic word. The difference between the two groups of languages is that, while in EP and Italian the prominent element of the compound prosodic word is its rightmost prosodic word, in Dutch and English, it is its leftmost prosodic word. These languages would thus behave in the same way in what concerns the phonological phrase, a righthanded prosodic domain that includes non-compound words as well as some compounds, namely those that display the weak-strong pattern in Germanic languages and non-lexicalized syntactic compounds in EP and Italian.

Not only the phonological evidence provided by EP, but also the stress pattern of Germanic languages compounds and of Dutch derived words including more than one prosodic word, as well as data from Baule reported in Leben and Ahoua (1997), strongly indicate that some prosodic words are grouped together under a prosodic constituent that does not behave like the phonological phrase, and which we have proposed to be a compound prosodic word. This suggests, therefore, a refinement in the analyses that treat all compounds as directly dominated by the phonological phrase (cf. Nespor and Ralli 1996 and Nespor 1999a, for a wide variety of languages, and Pereira 1999, for EP).

We believe that the compound prosodic word analysis proposed in this book may also account for several other facts that were previously explained resorting to the clitic group, like stress assignment within compounds in Hungarian (cf. Vogel 1990) and penultimate schwa deletion in French (cf. Hannahs 1995a), among many others.

5. Our findings and phonological theory

The phonological data described here has argued for analyses that make crucial use of compound and adjunction structures, in a violation of the Strict Layer Hypothesis (SLH). This may be seen in the tree presented in (483), which illustrates the prosodization of some of the structures investigated.³⁸⁴ Thus, EP provides additional arguments in favor of the weakening of SLH, as well as against the clitic group as a constituent of the prosodic hierarchy (cf., among others, Selkirk 1996; Booij 1996a; Peperkamp 1997a; see also chapter 1, sections 4–6).

U

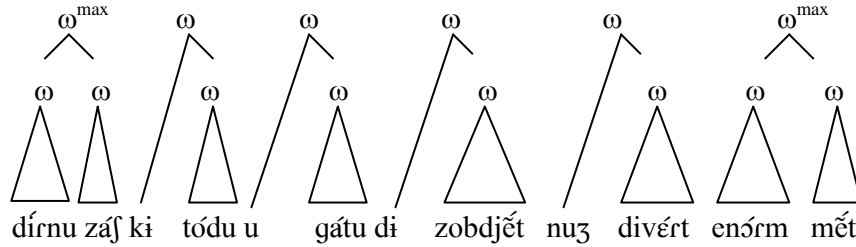
...

...

...

...

...



We believe that this work reinforces contemporary theories of prosodic phonology. In fact, the basic concepts and tools developed within this framework, which were broadly presented in chapter 1, proved very successful in accounting for many facts of EP. We hope, furthermore, that our analysis of this particular language may contribute to the understanding of yet not fully comprehended issues, such as which morphosyntactic structures are mapped onto which prosodic domains, what is the exact information required for the construction of the prosodic word domain, what are the structural restrictions imposed on the architecture of prosodic trees, or what are the sources and limits of crosslinguistic variation in the prosodization of morphosyntactic material.

With regard to phonological theory in general, we highlight one major aspect that follows from the present investigation. The organization of grammar as assumed here has proved extremely suitable to handle an important array of data. We have seen that, besides the importance of taking the prosodic word into consideration in the phonological analysis of EP, it is essential to acknowledge the separation of the lexical and postlexical components in order to establish important generalizations in the language. Along the lines of proposals within lexical phonology, reviewed in chapter 1, section 3, we have assumed that lexical and postlexical phenomena are characterized by well-defined and complementary sets of properties that allow us to assess the locus in the grammar where they are active. The

identification of the phonological events that belong to each level in EP was systematically carried out in chapter 3. Subsequent chapters showed the productivity of the proposed typology. This is particularly evident when pronominal clitics and affixes are compared (chapter 4). Whereas affixes interact with general phonological phenomena that operate within the lexical component, such as word stress assignment, nasal segment resolution, heterosyllabic /e/-centralization, glide insertion, among others, pronominal clitics never do. These facts, together with data from other components of the grammar indicating that clitics exhibit a syntactic behavior that is not compatible with their hypothetical affix status, were seen to constitute firm evidence for the postlexical attachment of pronominal clitics to their host.

The distinction between lexical and postlexical phonological phenomena was also seen to have implications on the identification of the prosodic structure assigned to morphosyntactic material. The comparison between pronominal cliticization in EP and in Standard Italian is particularly suitable for demonstrating the relevance of such a distinction. In both languages postverbal pronominal clitics do not affect the location of word stress in the verbal host. By contrast, in other varieties of Italian, such as Neapolitan and Lucanian, clitics either receive secondary stress or bear the primary stress of the host-clitic combination. In order to account for this asymmetry, Peperkamp (1997a) proposes that clitic pronouns in these varieties display distinct prosodizations. With respect to Standard Italian, postverbal clitics are assumed to be attached at the level of the phonological phrase, unlike in Lucanian and Neapolitan, where they are incorporated into and adjoined to the prosodic word, respectively. Standard Italian clitics, therefore, fall outside the domain of word stress assignment, and thus the lack of interaction between clitics and word stress is derived (see also chapter 1, section 6.3). Although EP postverbal clitic pronouns behave similar to Standard Italian pronouns as far as word stress is concerned, we have provided evidence that these elements are incorporated into the verb's prosodic word in EP (chapter 5, section 4.3.1). We have further shown that word stress in this language has the properties of a lexical phenomenon (chapter 3, section 2). Under our account, EP enclitics do not affect stress location in the host because word stress does not operate postlexically, and enclitics are combined with their host only at the postlexical level. These facts demonstrate that stress location is not a reliable diagnostic for the postlexical prosodization of clitics in EP. Consequently, with respect to Standard Italian, we believe it is necessary to demonstrate that word stress is active postlexically before it can be used as crucial evidence for the postlexical prosodization of pronominal clitics in this language.

Another area where the distinction between lexical and postlexical components becomes apparent concerns the behavior of the prosodic word domain itself. In chapter 3, we have offered arguments for the existence of the prosodic word domain already within the lexical component, thus extending to EP proposals already made on the basis of data from other languages (cf., among others, Booij 1988; Booij and Lieber 1993; Nespor 1990; Peperkamp 1997a; see also chapter 1, section 3.2). Nevertheless, additional phonological evidence was seen to indicate that a domain with similar characteristics, but with a few modifications, is obtained postlexically. We have accounted for these facts with the assumption that the complete construction of the prosodic word domain is accomplished in two stages. At the lexical component, morphological units that are lexically combined are arranged around the prosodic word domain: the units that bear their own stress form their own prosodic word, namely, stems plus suffixes that do not form independent stress domains, as well as affixes that constitute independent stress domains; and stressless prefixes are adjoined to the prosodic word that contains their base. As for clitics, they remain undominated by the prosodic word node in the lexicon, given that they are only combined with their host postlexically. In the postlexical component, some adjustments are required in order to allow not only for the prosodization of stressless units that are combined with their host postlexically (clitics), but also for the prosodization, within a single prosodic word, of sequences of prosodic words containing material that is dominated by a single syntactic terminal node.

Phonological processes and resyllabification in EP also provide evidence for the separation of lexical and postlexical components. We have seen in chapter 3 that processes that have the properties of lexical phenomena, such as the lowering of word final vowels in syllables closed by /r/, and of vowels in syllables closed by /l/, require the lexical syllabification of word final consonants (cf. chapter 3, section 8). Similarly, in the approach of Morales-Front and Holt (1997), adopted in this book, lateral and nasal segments assimilate the place feature of the following word final fricative (the plural morpheme /s/), which presupposes the lexical attachment and (at least partial) feature specification of word final fricatives (cf. chapter 3, section 5). Furthermore, none of these processes are affected by postlexical resyllabification which may cause the final consonants /r/, /l/, and /s/ to become syllable initial if followed, within the intonational phrase domain, by a word starting with a vowel or a glide. EP behaves like other Romance languages in this respect (e.g. Nespor and Vogel 1986; Booij 1988; Hall

1994; Peperkamp 1997a), and thus adds to the evidence that has been put forward against Rice's (1989) proposal that word final consonants are not lexically syllabified (see chapter 1, section 7 for a review on this issue).

6. Our findings and EP phonology

We think that our work gives a contribution to the understanding of EP phonology as well. This follows partially from the fact that, to the best of our knowledge, this is the first time that the prosodic word of EP is studied. We have also contributed to a more comprehensive and systematic separation of the lexical and postlexical phonological phenomena of EP. In fact, for a number of phonological phenomena investigated, no classification was previously proposed. In those cases where there had been such proposals, we have been able, in general, to add novel arguments in support of a classification that which was not always coincident with previous accounts. We believe, furthermore, to have contributed to a better understanding of phenomena that are more familiar aspects of the language. We highlight, in particular, our analysis of non-back final vowel deletion, of the realization of word initial vowels, of the realization of stressless vowels in word final syllables closed by sonorants, of non-primary stress assignment, and of vowel reduction and other reduction phenomena (see chapters 3, 5 and 7).

In the realm of the prosodic phonology of EP, the bulk of evidence for the prosodic word domain manifestly contrasts with the paucity of phonological phenomena associated with the prosodic domains that occupy adjacent positions in the prosodic hierarchy, namely, the phonological phrase and the foot. A similar contrast is noticed by Frota (2000) between the phonological phrase and the (phonologically much richer) intonational phrase domain. The following list includes most of the evidence identified so far for the syllable, the foot and the prosodic word domain.

(484) *The Syllable:*

- (i) a domain referred to by many phonological processes – e.g. coda *l*-velarization, fricative and nasal consonants feature specification, vowel nasalization, word internal *r*-strengthening (cf. Mateus and d'Andrade 2000, among others);
- (ii) a domain that defines segmental (co-)occurrence restrictions – e.g. syllable initial consonant clusters are allowed, but only if the first consonant is an obstruent and the second one is /r/ or /l/; final consonant clusters are usually not allowed; among all consonants

of the language, only /r/, /l/, the coronal fricative consonant, and nasal segments may close a syllable (cf., among others, Vigário and Falé 1994; Mateus and d'Andrade 2000);

(iii) possibly, the unit to which the tone units that make up the intonation contour are associated (Frota 2000);

(iv) a domain with respect to which EP speakers have, in general, clear intuitions.

The Foot:

No phenomena have been argued to clearly refer to the foot in EP.³⁸⁵

The Prosodic Word:

(i) a domain referred to by many phonological processes – e.g. lowering of stressless vowels in final syllables closed by /r/; initial /r/-strengthening; initial vowel feature specification; V2 semivocalization; final non-back vowel deletion; final central and round vowel deletion; syllable degemination (cf. chapter 3);

(ii) a domain for phonotactic constraints – e.g. prosodic words in EP do not start with [ʎ], [ɲ], [r], and [i] (cf. chapter 5, section 2);

(iii) the domain of several stress-related phenomena – e.g. the domain that includes units with a primary stress; the domain that defines the context for the assignment of secondary initial stress and initial emphatic stress; the domain referred to by focal stress assignment (see, in particular, chapter 5, sections 2 and 4.3.2, and chapter 6, section 3, respectively);

(iv) a domain relevant for tonal association – e.g. the presence of pitch accents signals a prosodic word, as well as the head of a compound prosodic word; the prosodic word defines the constituent whose right-boundary is assigned the intonational phrase H-initial tone (see, in particular, chapter 3, section 18, and chapter 6, section 3.1);

(v) a domain referred to by constituent deletion processes – e.g. deletion under identity and clipping (cf. chapter 6, section 4).

Frota (2000: 2.3) has interpreted the difference in phonological visibility between ϕ and I as resulting from the “strength” of each prosodic domain, the former being weak and the latter being strong. Although it is not obvious how the concept of strength of prosodic domains can be integrated into

prosodic phonology theory, at the observational level it is clear that the same type of distinction sets apart, on the one hand, the syllable and the prosodic word, which display the phonology of strong domains, and, on the other hand, the foot and the phonological phrase, which behave like weak domains. Interestingly, Frota (2000: 102) puts forward the hypothesis that strong categories, but not weak ones, favor compound domain formations, as verified at the level of the intonational phrase in EP (see chapter 1, section 5 for a review of Frota's arguments in favor of I-compounding). As we have shown in chapter 6, this prediction is borne out for the prosodic word domain as well.

Besides the aspects mentioned in Frota (2000), it is conceivable that the difference in functionality between prosodic domains may be behind some crosslinguistic variation regarding the selection of different prosodic constituents for particular phenomena. We may illustrate this point with the crosslinguistic variation found in the processes of clipping and deletion under identity with respect to the prosodic domains they refer to.

In several languages, clipping crucially refers to the foot domain (e.g. Mester 1990 for Japanese, Prieto 1992 for Spanish, Cabré 1993 and Cabré and Kenstowicz 1995 for Catalan, Thornton 1996 for Italian – see chapter 1, section 6.1 for a review). In EP, we have shown that clipping exclusively refers to the prosodic word (see chapter 6, section 4). Language selection of one prosodic domain or the other should not be attributed to, say, morphological differences between languages. This is demonstrated by the comparison between EP and BP. Although these varieties of Portuguese show in general identical morphology, in the latter clipping appears to be based on the foot, similar to other Romance languages, but unlike EP. This is shown by the BP data in (485) (from Araújo 2001; boldface signals the syllable bearing secondary stress, and capitals signal the word stressed syllable).

(485) <i>dinoSSAUro</i>	>	<i>DIno</i>	<i>refrigeRANte</i>	>	<i>reFRI</i>
'dinosaur'			'juice'		
<i>retarDAdo</i>	>	<i>REta</i>	<i>refrigeraDOR</i>	>	<i>REfri</i>
'retarded'			'refrigerator'		
<i>delegaCia</i>	>	<i>deLE</i>	<i>profisioNAL</i>	>	<i>proFI</i>
'police station'			'professional'		

Notice that stress location in the clipped form seems to correspond to the secondary stress of the input form. This suggests that clipping in BP reduces the input expression to the first foot.

The fact that BP has other phenomena that cue the foot domain is further suggestive of the strength of this domain in this variety, in contrast with EP. For example, the rhythmic alternation of strong-weak syllables is repeatedly referred to in the literature on BP (e.g. Brandão de Carvalho 1989; Lee 1995; Moraes 1997; Collischonn 1999). This pattern may be taken to correspond to the prominence relations established at the level of the foot. In addition to this, BP has phonological rules that are argued to refer to the foot. This is the case of Posttonic Vowel Neutralization, which is responsible for the neutralization of the opposition between [o] and [u] in posttonic position (cf. Wetzels 1992).³⁸⁶

Like clipping, deletion under identity in BP appears not to refer to the prosodic word, despite the fact that this domain is crucial for the process in other languages, such as Dutch, German, and EP (cf., respectively, Booij 1985, 1995, Wiese 1996, and chapter 6, section 4). Indeed, the examples in (486), accepted as possible in Schwindt (2000), citing Moreno (1997), show expressions containing prefixes which have been argued not to bear word-primary stress in BP (cf. Schwindt 2000).

- (486) *exporTAR e imporTAR* > *ex e importar* 'import and export'
incluSive e excluSive > *in e exclusive* 'inclusive and exclusive'

We should remark that EP allows the coordination of the very same expressions, but it does not allow deletion under identity, because the relevant coordinated units are not composed of more than one prosodic word. The data in (486) lead Schwindt (2000) to claim that deletion under identity may be used as a criterion for the identification of synchronic affixes, but not of prosodic words. It appears, therefore, that the prosodic word is not the domain referred to by this process. This may suggest that the prosodic word in BP is a weak domain, in contrast with the foot.

We may further notice that, at least from the work of Schwindt (2000) and Bisol (2000), it is not phonologically clear what the precise prosodic relationship between clitics and adjacent prosodic words is in BP. This contrasts sharply with the facts of EP. Thus, the prosodic word in BP is possibly a domain weaker than in EP.³⁸⁷

In conclusion, the facts described in the preceding paragraphs suggest that typologically related languages and varieties may exhibit clear differences in the role played by a given prosodic domain in the phonology of the language. A more comprehensive comparison between EP and other languages is, nevertheless, required in order to fully understand the actual

relevance of these observations.³⁸⁸ This is certainly a topic worth exploring in future work.

7. Our findings and language acquisition

Besides showing sound evidence for the prosodic word domain, EP provides also very clear evidence for the lack of prosodic word status of clitic function words. Not only are clitic words stressless and undergo vowel reduction, but in addition, sandhi processes and other phonological phenomena of the language seem to systematically distinguish clitic words from full prosodic words (see, in particular, chapter 5, section 4, and chapter 7). This distinction is so remarkable that it is likely to play some, for instance, in language acquisition. This idea is, in fact, pursued in Selkirk (1996), Morgan, Shi, and Allopenna (1996), and Peters and Strömquist (1996), among others, who suggest that the establishment of the distinction between lexical and grammatical words can be performed at least partially on the basis of the phonological properties associated with both categories of words. That is, the systematic phonological distinction between (clitic) function words and lexical words may trigger the acquisition of the basic syntactic distinction between lexical and functional categories.

There are other areas where language acquisition may take advantage of prosodic phonological information. Not only EP phonology signals the distinction between lexical and grammatical words, but it also allows for the identification of morphosyntactic constituents. For example, prefixes and proclitics are prosodized as adjuncts to an adjacent prosodic word, and therefore may be distinguished from elements that belong to a stem/word on the basis of their phonological behavior; enclitics often violate the three syllable stress window, and thus can be differentiated from suffixes, on the one hand, and be set apart from the verbal host itself, on the other hand; and the presence of more than one word stress, together with other phonological properties of prosodic words, identifies the complex internal structure of morphological compounds, as well as of other syntactic X^0 -level expressions. It is thus conceivable that phonological information may be used in the acquisition of these morphosyntactic categories.

Prosodic information may also provide the basis for inferring other aspects of syntactic structure. For example, Mazuca (1996) and Nespor, Guasti, and Christophe (1996) propose that the setting of the syntactic parameter that determines the language directionality may be established on the basis of the prominence relations that hold within prosodic phrasal do-

mains. Along the lines of much previous work in prosodic phonology (e.g. Nespor and Vogel 1986), Nespor, Guasti, and Christophe (1996) observe that languages show a correlation between right/left syntactic branching and ϕ -level final/initial prominence, respectively. Given this correlation, the hypothesis is put forth that the setting of this syntactic parameter is performed on the basis of the acquisition of the prominence relations at the level of ϕ – which is most likely to be very early, since it has been shown that newborns are sensitive to stress patterns. This is an extremely appealing hypothesis because, according to it, the child could be able to set a core syntactic parameter before (s)he acquires any syntactic knowledge and before segmenting.³⁸⁹ Interestingly, our compound prosodic word hypothesis for derived and compound words displaying strong-weak pattern in English and Dutch becomes crucial in the context of Nespor, Guasti, and Christophe's proposal. Indeed, if such compound and derived words were to be analyzed as forming ϕ 's, and not compound prosodic words, then the ϕ -domain in these languages would have to allow for both initial and final prominence. In such a case, the child would be exposed to contradictory information which could impinge on the setting of the syntactic parameter of direction of recursion. By contrast, if the two stress patterns are associated with different prosodic configurations, this problem is solved.

In the same way phonological information at the ϕ -level may trigger the setting of the word-order parameter, other aspects of syntactic structure may be conjectured to be deduced from prosodic constituency at the level of the prosodic word. The prosodic differences displayed by compound words and syntactic phrases in languages like English, Korean and Japanese have been suggested to possibly constitute cues for the child to differentiate between these syntactic objects (cf. Venditti, Jun, and Beckman 1996). In the same vein, given the phonological evidence that EP provides for compound prosodic words, as well as for the specific prosodization of enclitic pronouns, it can be hypothesized that the child acquiring this language may infer on the basis of prosodic information the syntactic nature of complex expressions that are grouped together under a syntactic terminal node.

To sum up, the contribution to prosodic phonology made in this work may inform future research on the role that phonological properties play in language acquisition.

8. Final word

In this book we have provided a detailed description of EP data, together with a clear presentation of our findings and their consequences to the theory of Prosodic Phonology. We hope this may stimulate the inclusion of European Portuguese in future crosslinguistic discussions about the topics studied in this book. In the same way we feel eager to proceed with the study and follow the path of the questions left unanswered, we hope that this work may provide motivation for others to embark on the adventure of seeking to understand more about the phonology of (this) language.

³⁸² We believe that this proposal may also be extended to BP. The importance of morphological information for the prosodization of prefixes may explain why, according to Schwindt (2000), synchronic (non-lexicalized) prefixes in BP are prosodized similar to EP (cf. Vigário 1999a; see also section 5.2.2). As for the prosodization of proclitics in BP, the data in Schwindt (2000) and Bisol (2000) clearly indicate that, like prefixes, proclitics are not incorporated into the prosodic word that dominates their base/host. Nevertheless, the identification of the type of phonological host and the prosodic relation between proclitics and host remains unclear. These issues are not systematically addressed in Schwindt's work because he adopts the view that clitics are part of the clitic group. Bisol (2000) is less categorical in this regard, admitting that the clitic group may be reinterpreted as a postlexical prosodic word.

³⁸³ We should observe, nonetheless, that word-level stress has consequences in the definition of the (lexical) prosodic word in EP similar to phonotactic constraints in Dutch and German. That is why we have proposed in section 6.4.2 that not only morphosyntactic but also purely phonological information is required for the construction of this prosodic domain in EP.

³⁸⁴ We assume that resyllabification induces prosodic word restructuring, along the lines of Peperkamp (1997a) (see section 1.6).

³⁸⁵ To the best of our knowledge, EP does not present any phenomenon that undoubtedly refers to the foot. With respect to word stress, its location appears to be obtained to a great extent from lexical and morphological information (see Mateus 1983 and the literature reviewed in chapter 2 and section 3.1). Except for word-stress – if at all –, it is not clear what role this constituent may play in the phonology of the language (see also section 5.3.3.2, for a discussion on secondary stress in EP).

³⁸⁶ In EP this type of neutralization affects other vowels, as well, and applies to unstressed vowels within the word, regardless of their location with respect to word stress.

³⁸⁷ Although compounds are in general not referred to display a phonological behaviour distinct from phonological phrases in BP, Lee (1995: 3.2) suggests that stress-clash resolution does distinguish between lexical compounds and sequences of words combined postlexically, in that it may apply to the former but not to the latter. He suggests, furthermore, that stress-clash resolution in BP is a lexical phenomenon. However, clash resolution in this variety is reported in other studies to apply optionally within the phonological phrase (cf., among others, Major 1985: 4.2; Abousalh 1997). According to L. Bisol (personal communication), at least in some varieties of BP stress clash resolution is obligatory within compounds but optional within phonological phrases.

³⁸⁸ Kleinhenz (1997) proposes that some languages are word-based and some are phrase-based, in the sense that the prosodic phonology of languages is organized on the basis of the prosodic word or the phonological phrase, respectively. EP is

given as an example of a word-based language, whereas French and Chi Mwiini represent phrase-based languages. German and Dutch are proposed to be intermediate stage languages. The argument for the classification of EP as word-based is founded on Frota's (1996) observation that ϕ is in general invisible in EP. The data gathered in this book can now be added in order to support the view that EP phonology extensively refers to the prosodic word. It should be noticed, nevertheless, that phonological phenomena that refer to ϕ , although rare, do exist in EP (see chapter 1, §1.1.2, and note 9 below).

³⁸⁹ We have seen that there is language variation in the role that prosodic constituents play in particular phonological systems. What the consequences of this type of crosslinguistic variation for language acquisition may be is too soon to tell. We may observe, nevertheless, that work on EP has shown that ϕ is a weak prosodic domain (e.g. no sandhi phenomena refer to this constituent, it is often not tonally marked, and frequently ϕ -level prominence is not perceived – cf. Frota 2000; Vigário 1998a). Despite this, children do not seem to have any difficulty in acquiring the word order of the language. Interestingly, the (scarce) evidence that has been gathered for ϕ in EP consist of phenomena that somehow refer to the stress or the head of this constituent (see section 1.1). Whether this may indicate that there is a limit to phonological constituents' weakness in those domains that play a major role in the acquisition of grammatical information, is a question we leave open.

Notes

1. Prosodic phonology is first developed as a reaction against Chomsky and Halle's (1968) conception that the syntactic surface structure and the phonological representation are related via a set of readjustment rules, which introduce boundary symbols into the surface syntactic structure. According to this approach, phonological rules refer to such boundary symbols in their structural descriptions, rather than directly to syntactic bracketing (see, for example, Selkirk 1980; Hayes 1989: 2.1.2).
2. Although the syllable and the foot are not treated in earlier work (e.g. Selkirk 1980; Nespor and Vogel 1982), in Nespor and Vogel (1983, 1986) the prosodic hierarchy is already assumed to include these constituents. Itô (1986) and Zec (1988), cited in Inkelas (1990), propose, in addition to these constituents, the inclusion of the mora in the prosodic hierarchy. For Hayes (1989) the lowest constituent of prosodic hierarchy is the phonological word. Inkelas (1990), in turn, proposes the replacement of the mora, the syllable and the foot (considered to belong to the metrical structure) by other sublexical prosodic constituents, built with reference to the morphological structure. The main motivation for this view is to prevent lexical rules from referring directly to morphological structure. Selkirk (1986) also questions the status of the syllable and the foot as prosodic domains, although in later work (e.g. Selkirk 1996), the prosodic hierarchy is assumed to include these constituents.
3. In this book, we use the term prosodic word, instead of phonological word with which it is synonymous. We should point out that the latter was already at use before the development of prosodic phonology (see, among others, Câmara 1953, 1972; Dixon 1977).
4. Since the universal inventory and definition of prosodic domains is ultimately an empirical question, it is natural that some controversy in the literature is found. Thus, for example, on the basis of Greek data Condoravdi (1990) proposes a new prosodic category above the ϕ -level (see section 5 on an alternative proposal).
5. Languages like German and Dutch also show this type of behavior with respect to the phonological phrase: while there are no segmental rules that refer to this prosodic domain, there are other types of phonological phenomena that cue the ϕ -phrase (see Kleinhenz 1997 for relevant data and discussion).
6. The hierarchical nature of phonological representations is also a major departure from SPE linear conception, and it originates, in particular, with Liberman and Prince's (1977) proposal on English word and phrasal stress. Note that, under the latter approach, the hierarchical grouping of units is

- binary, and thus it creates a structure with unlimited depth. This property distinguishes metrical trees from prosodic trees, since the latter are flatter and their depth is limited (see, for example, Pierrehumbert and Beckman 1988: 6.1).
7. Early examples of prosodic structures that do not obey strict layering can be found in Nespor and Vogel (1983) and Nespor (1986). For example, in Nespor and Vogel (1983) restructuring at the ϕ -level results in the creation of a ϕ' -node, defined in the following terms: a nonbranching ϕ which is the first complement of X on its recursive side loses its label and is joined to the ϕ containing X , under a new node labeled ϕ' (where X is a lexical head) (Nespor and Vogel 1983: 126).
 8. “Lex” stands for a lexical head and “Lex^{max}” stands for the maximal projection of a lexical head. This formulation excludes maximal projections of functional categories from playing a role in the syntax-phonology mapping (see Truckenbrodt 1999 for a review and discussion).
 9. Under Frota’s (2000) view, the weight conditions on I and ϕ -phrases play a role in the *basic* formation of these constituents, so that the mechanism of restructuring (Nespor and Vogel 1983, 1986) is no longer required. Notice, further, that the application of the weight conditions is constrained by the limits defined in (5) and (8).
 10. The prosodization of clitic function words is not discussed in Frota (2000), who adopts basically the analysis proposed in Vigário (1999a), and developed in this book. Specifically, proclitic function words are assumed not to incorporate into the host prosodic words, and thus they count for the purposes of ϕ -branching.
 11. Although we do not have clear evidence for the foot domain in EP, the evidence for this domain in other languages argues for its inclusion in the prosodic tree. Since the prosodic tree is assumed to be universal, the foot should integrate EP prosodic trees as well. We should notice, additionally, that in example (11) we have avoided clitic elements, since one of the central questions addressed in this book concerns precisely the way these elements are prosodized. Prominence relations within prosodic constituents are not indicated (see §5).
 12. Frota’s contention is that this result is a derived one, since she claims that only in languages where focus is morphosyntactically marked is the prosodic structure directly affected by focus. Since EP has no such morphosyntactic properties, focus information does not affect prosodic phrasing in this language.
 13. The following typology of prosodic processes may be drawn according to the way that phonological processes apply with respect to prosodic constituents (cf. Selkirk 1980; Nespor and Vogel 1986: 15):

Types of prosodic processes

- *Domain-span*: processes that apply *within* a given prosodic domain
- *Domain-limit*: processes that apply at the *edge* of a given prosodic domain
- *Domain-juncture*: processes that apply at the *edge* of a given prosodic domain, *within* a larger prosodic domain.

14. As Anderson (1992: 7.1) points out there are at least three notions of lexicon: (i) the locus of idiosyncrasy in language; (ii) the collection of all the items that belong to open or major word classes; and (iii) the scope of Lexical Phonology, as opposed to the domains of other processes characterized as postlexical. In this book we use the term *lexical entry* as the point where idiosyncratic information of individual items is stored, and the terms *lexical component* or *lexical level* to refer to the point in the grammar where lexical processes apply.
15. Hannahs (1995a: 5.3) suggests that the need for lexical levels can be correlated with the history of languages. Thus, languages where important parts of the lexicon are of different origins have been described as requiring the partitioning of the lexicon (as in the case of English, or Malayalam), whereas no such partition is necessary for languages with a lexicon of essentially a single origin (as in the case of French, or Italian – cf. Vogel 1991). As for the ordered conception of lexical levels, see Booij (1994: section 5), which also contains a review of important arguments against such a view.
16. The elimination of level ordering is not incompatible with the conception that lexical phonology precedes postlexical phonology, since the latter follows in a principled way from the organization of the grammar that is assumed (e.g. Booij 1996b).
17. As we will show in chapter 3, a number of lexical processes in EP are not structure preserving (e.g. vowel reduction – section 3; processes that create nasal vowels and glides – section 5; /e/-centralization – section 6; initial /r/-strengthening – section 9).
18. Precompiled phrasal rules have been proposed to account for similar effects in several other languages as well (e.g. Italian – cf. Nespor 1990, Peperkamp 1997a; Galician – cf. Fernández Rei 2002). Some facts of EP are also assumed to follow from precompilation in Vigário (1998b) and chapter 4, section 4.2.2 of this book.
19. For Kaisse (1990), there are two types of rules that may apply across words: P1 rules, which share a number of properties with lexical rules; and P2 rules, which have the properties considered below to characterize postlexical rules. According to Kaisse, both types of rules apply postlexically, P1 rules preceding P2 rules. Notice, nevertheless, that the existence of P1 rules does not preclude precompilation, as Kaisse remarks. In fact, if a rule that applies between words precedes (other) lexical processes, as is the case of vowel shortening in Kimatuumbi (Hayes 1990), then that rule cannot be postlexical – it must be a lexical one. In this book, we follow Hayes' proposal that if a rule

- applying between words has the properties of a lexical process, then it is a lexical process. In fact, as far as we can see, EP does not provide evidence for a split in the postlexical component.
20. Hayes suggests, additionally, that precompiled rules are not sensitive to speech rate. However, Nespor (1990) holds that they may be sensitive to rate if the frame that defines the context for the application of precompiled rules includes a prosodic domain, as in the case of *Troncamento*.
 21. Sensitivity to speech rate may be, at least partially, a consequence of the sensitivity of prosodic constituents to speech rate: if a process applies within a given prosodic domain, and if that domain may be smaller or larger depending on the speech rate (see §7), then sensitivity of that process to speech rate may result from the variability of the prosodic domain limits.
 22. A more or less widespread idea within lexical phonology is that function words, which do not belong to the major categories (N, V, and A), and do not enter in morphological operations, unlike affixes, are not present in the lexical component, and therefore they never undergo lexical rules (see the observations in Kaisse and Shaw 1985: 9, and, for an explicit claim along these lines, Inkelas 1990: 241). We show in chapter 5, section 4.2 that some EP facts contradict this view.
 23. See also sections 6.3 on the prosodization of clitics and 7 on resyllabification.
 24. Frota (2000) adopts Ladd's (1992, 1996) Compound Domain Hypothesis but she suggests that, at least for the intonational phrase level in EP, compound domains should be constrained in order to allow only for binary recursion.
 25. As shown in chapter 6, EP presents evidence for prosodic compounding at the prosodic word level as well. The analysis proposed there is extended to languages like English, Dutch, Italian, French and Baule (see also chapter 8).
 26. For Ladd (1992) only two conditions are needed to restrict the prosodic structure: the Ranking condition (which corresponds to Selkirk's Layeredness constraint), and the Uniformity condition, which is stated in (i).
 - (i) Uniformity condition
In a prosodic tree, all nodes immediately dominated by a given node N must be of uniform rank.
 27. Compound domains have been suggested to be the solution for some of the cases reported in the literature to weaken prosodic phonology theory and its underlying assumption that phonological processes do not refer directly to syntactic information. For example, under his Compound Domain Hypothesis, Ladd (1992) proposes to analyze Makonde's tone sandhi domain (presented in Odden 1990) as a compound domain, thus avoiding an approach with direct reference to syntactic information; and Frota (2000) suggests that compounding could also account for Greek data that led Condoravdi (1990) to postulate the existence of an additional prosodic domain ('Z') between the phonological phrase and the intonational phrase.

28. French has been argued to have a stress domain larger than the word (cf. Pulgram 1970; Dell 1984, cited in Van der Hulst, Hendriks, and Van de Weijer 1999; Van der Leeuw 1997). On the contrary, Ladd (1996) claims that tonal distribution in French, as described by Dell (1984), can only be accounted for if it is assumed that the last full vowel of a prosodic word is metrically strong. Roca (1999) also analyses French, like other Romance languages, with a primary stress assigned at the word-level. This is, in fact, what is predicted by prosodic phonology theory – since it is assumed that each prosodic domain has a prominent element – independently of the possibility that phrasal stress might be more clearly perceived than word stress in some languages. Besides this, Hannahs (1995a, 1995b) presents segmental evidence for the prosodic word in French (e.g. glide formation).
29. See Booij (1999), who shows that the relevant generalizations are constraints on prosodic rather than morphological constituents.
30. In fact, we will show that in intonational languages, like EP, several types of tonal event also cue the prosodic word (see chapter 3, section 18).
31. Hall (1999a) notes, nonetheless, that there is one suffix in German (*-chen*) that can be deleted under identity, but does not show the behavior expected of a prosodic word, since it has a schwa as its only vowel.
32. Booij (1985) proposes in addition that reduction is restricted to occur only at the edges of phonological phrases, and that the two phonological phrases including the identical prosodic words must be adjacent.
33. Notice that this type of operation is hypothesized in Zwicky (1977: 3) to be ruled out by a “Rule Immunity”, which states that “proper parts of words do not undergo rules of deletion under identity”.
34. In chapter 5, section 2 we discuss in more detail the relevance of defining a minimal word size in European Portuguese, comparing it in particular with Italian. In fact, this language is similar to Portuguese in this respect, but has been claimed to have minimality requirements on the basis of the low frequency of monosyllabic words with open syllables (cf. Thornton 1996 and Peperkamp 1997a; see also Vogel 1993, and, for a different view, Bafile 1997).
35. Wennerstrom (1993) considers that the presence of focus is also a test for the identification of prosodic words (a test also used in Hall 1999a for German). Raffelsiefen (1999a) argues against this position, since the possibility of focus may not correlate with other properties that can be established to identify English prosodic words. The notion of focus used by Wennerstrom seems parallel to the one we called initial emphatic stress in EP (see chapter 3, section 17). As we will see in chapter 5, in EP emphatic stress (but not focal stress) may be assigned both to *stressless* prefixes and to proclitics, which are independently shown not to form autonomous prosodic words.
36. By contrast, Hayes (1989: 207) assumes that the prosodic word is always at least as large as the grammatical word. This view does not account for the existence of languages like Dutch where some suffixes behave like independent prosodic words.

37. In Nespor and Vogel's definition of the prosodic word domain of Type II languages, there is an additional clause that we have not included in (32), stating that unattached elements form independent prosodic words. This is meant to allow for elements that do not include stems, like function words, to form their own prosodic word. Under an analysis where these elements, being typically clitics, are not prosodic words, it is no longer necessary to include this information in the definition of the PW domain (see next section, on the prosodization of clitics).
38. See Peperkamp (1997a: 5.4.1.2) for the discussion of an alternative analysis of Neapolitan facts whereby the verb and the clitic each form an independent prosodic word.
39. The prosodization of Standard Italian enclitics is discussed in chapter 8 in the light of our proposal on the prosodization of EP enclitics.
40. In yet other proposals, such as Berendsen (1986), clitics are conceived as floating elements whose prosodization may freely vary. For example, in Dutch they would optionally incorporate into the prosodic word, the phonological phrase, or both.
41. The arguments for the proposal that in the most frequent case enclitics in German are incorporated into the preceding prosodic word are weaker in the sense that the data is consistent with, but does not force, such an analysis. This proposal contrasts with Kleinhenz's (1996) analysis of German enclitics, according to which these elements are adjoined to the preceding prosodic word.
42. In a theory such as Inkelas' (1990), the direction of cliticization follows from the prosodic subcategorization frame of clitics. In this case, therefore, the direction of cliticization is always a property of individual clitics. Such a theory does not account for the optionality in the direction of cliticization found, for example, in Dutch (see also Van der Leeuw 1997).
43. This is proposed to be the case of all European languages surveyed in Nespor (1999b), including, besides Greek and Italian, Dutch, Icelandic, Spanish, French, Romanian, and Portuguese.
44. Nespor (1999b) claims that word+word compounds are always mapped onto two prosodic words, a hypothesis supported by many Romance, Germanic, and Slavic languages.
45. Other languages where stem+word compounds are argued to be mapped onto two prosodic words are Catalan, Romanian, and Serbo-Croatian (Nespor 1999b).
46. In order to derive the latter possibility, Peperkamp (1997a: 127) proposes a constraint that allows for familiar compounds to be mapped onto a single prosodic word.
47. See also Di Sciullo and Williams (1987), who propose that noun+noun compounds in French are in fact listed phrases, that is, lexicalized syntactic phrases.

48. There are, nevertheless, two such suffixes that do not form independent prosodic words in the language.
49. As shown in chapters 5 and 6, section 5 of this book, EP, like other Romance languages, patterns differently from Dutch and (most varieties of) German in that the segmental make-up of suffixes does not correlate with the way these units are prosodized. For instance, the great majority of consonant-initial suffixes are incorporated into the prosodic word that dominates their morphological base.
50. Notice, nevertheless, that Raffelsiefen (1999a: 143) shows that unproductive prefixes in English and in German can form independent prosodic words.
51. Peperkamp (1997a: 30) points out that there seems to be a correlation between the (im)possibility of phrasal resyllabification, and the type of syllables found in a given language. Romance languages, Turkish and Korean show simple syllable structure and allow for phrasal resyllabification, whereas Germanic languages show complex syllable structure and do not allow for phrasal resyllabification.
52. The term resyllabification may also be used to refer to the cyclic operation of syllabification that applies word-internally (e.g. Nespor and Vogel 1986: 106, Rice 1989). Here, we are only concerned with resyllabification across word-boundaries. The need for the distinction between (lexical) syllabification and (postlexical) resyllabification is demonstrated, for example, in Nespor and Vogel (1986: 68–72), on the basis of Spanish and French data.
53. In the examples, we omit the details concerning the syllable internal structure for ease of exposition.
54. The data provided by Dutch (or German, see for example, Wiese 1996; Kleinhenz 1996; Hall 1999a) involving vowel-initial clitics further show that, like in Romance languages, in Germanic languages (postlexical) resyllabification occurs, although it has a lower domain – the prosodic word.
55. This proposal is meant to account for resyllabification both between words and between prefixes (whether stressed or unstressed) and their morphological bases. In fact, the prosodic word boundary readjustment allows Peperkamp (1997a) to analyze Italian prefixes ending in a consonant and followed by a base starting with a vowel not necessarily as elements incorporated into the base prosodic word. As we have seen in section 6.2, the latter analysis is proposed in Nespor and Vogel (1986) for Italian prefixes ending in consonant, regardless of the presence of a context for resyllabification.
56. Researchers are not always aware of the importance of isolating this variable. In EP, the observation of more than one dialect in the description of a given process is possibly the most common practice. In other languages, such as German, we have seen that the prosodization of affixes may vary from dialect to dialect (cf. section 6.5). The same may happen with clitics (as is shown by Serbo-Croatian and Italian dialects – cf. section 6.3). Notice, nevertheless, that it can also happen that data from different dialects converge towards the

- same analysis, as is suggested to be the case of German clitics in Kleinhenz (1996). In any event, this source of variation should not be discarded.
57. An illustrative example of how this point may be crucial for a given analysis is provided by the account of the prosodization of pronominal clitics in Standard Italian, developed in Peperkamp (1997a) (see section 6.3). In this variety of Italian, the presence of enclitics does not affect the assignment of word stress: word stress falls only on the host, in the same place as when it occurs in isolation. This is taken to indicate that the enclitic occupies a position in the prosodic tree external to the prosodic word. Nevertheless, if word stress assignment is not operative postlexically, this behavior is in fact explained independently of the way postlexically combined elements are prosodized – the process is simply not active at the moment the relevant elements occur together. EP behaves similarly to Standard Italian in that pronominal clitics do not interact with word stress. Interestingly, this language provides evidence to choose between an analysis of this fact as following from the prosodization of enclitics, or as following from the non-application of word stress assignment within the postlexical component (see chapter 3, section 2, and chapters 5 and 8 for the discussion of this issue).
 58. In the examples of this chapter, as well as elsewhere in this book, a vowel bearing word-level stress is indicated with capital letters.
 59. For the distinction between the secondary stress that has the same general properties as the primary word stress (which is the one we are describing here) and the secondary stress that has rhythmic properties, see, among others, d'Andrade (1997) and Pereira (1999: chap.5), and references therein.
 60. For a number of phonological similarities and differences between the two varieties of Portuguese, see, for example, Parkinson (1988), Mateus and d'Andrade (2000) and Frota and Vigário (2000). Traditional grammars of Portuguese, such as Said Ali (1964) or Cunha and Cintra (1984), are also rich in references to some of the most salient differences between the two varieties.
 61. Despite the similarities among clitics also found in the European variety of Portuguese, we present in chapter 5, section 4.3 a number of facts that show that enclitics and proclitics behave differently from a phonological point of view with respect to their host prosodic word.
 62. As we will see in section 4.2. below, not all linguists accept that preverbal pronominal clitics are phonologically proclitic in EP.
 63. On the status of boundaries in SPE framework, see Chomsky and Halle (1968: chap.6) (see also chapter 1, section 2, note 1), where it is assumed that the boundaries signaled with = are stronger than formative boundaries (+), but weaker than word boundaries (#). For subsequent arguments against this approach to general juncture phenomena see, in particular, Selkirk (1980: section 3), and references therein, and for specific claims against the boundary =, see Siegel (1980).

64. In d'Andrade (1977), only the verbal system is considered in the discussion of word stress, although some remarks can be found on word stress in the nominal system (e.g. d'Andrade 1977: 25).
65. This is true at least of *z*-suffixes. In the case of *-mente*, Villalva (1992) assumes that the base this element attaches to is an adjective inflected in the feminine form that establishes a relation of agreement in gender with *mente* (which is diachronically related to the Latin word *mens*, *mentis* 'mind', from which the Portuguese feminine noun *mente* 'mind' is also derived). Villalva therefore proposes that this element is still analyzed as a word that belongs to the category N. We should add, nonetheless, that this analysis is not followed in latter work (Villalva 1994). The status of *z*-evaluative suffixes is not so clear, since, although Villalva considers these suffixes to be *words*, no category label is assigned to these elements, which form a left-headed root compound (see also Villalva 1994 for a different approach).
66. Notice that, by d'Andrade's convention (16), if two extrametrical elements exist in the same domain, only the second one remains extrametrical. Thus, unless different clitics are assumed to belong to different domains, in forms such as *faLa-r-se-te-Á* 'it will be told to you', the first clitic is not extrametrical and therefore the stress in *lar* may not be derived by the convention (48), which stipulates that the final vowel of the derivational stem receives stress when followed by an extrametrical vowel.
67. The following quotation synthesizes Van der Leeuw's reasoning: "...since Optimality Theory abandons rules, repair strategies or derivations, the complete phonological and prosodic structure of all candidates must be posited in one go. This means that there may be candidates with an extra primary stress on the stem and that we may freely insert prosodic boundaries. ...We do not have to derive a form with two prosodic words from one prosodic word, but instead we 'only' have to explain why in this case two prosodic words are better than just one." (Van der Leeuw 1997: 157).
68. It should be noticed, however, that the present indicative and the imperfect forms of *haver* do not necessarily (fully) coincide with the endings of verbs in future and conditional (see chapter 4, section 4.3, for an alternative proposal on the relation between the auxiliary *haver* and the last part of mesoclitic constructions).
69. This is also proposed in Vigário (1999a) and in Mateus and d'Andrade (2000).
70. For Pereira, the morphological domain for word stress is the "derivational stem" in the nominal system and the "word" in the verbal system. The derivational stem is defined as a constituent of the word that excludes inflectional affixes. This constituent *may* include one or more roots, as well as derivational affixes (cf. Pereira 1999: 129). Notice that this sort of definition appears to open the way for elements that do not contain a root to be analyzed as derivational stems.

71. This proposal is reminiscent to the idea suggested in Mateus and d'Andrade (2000) that the prosodic word status of *-mente* and *z*-evaluative suffixes is responsible for the postlexical location of the word-formation process that involves these suffixes (see section 4.3, below). It constitutes, in addition, a generalization of Nespor's (1999b) analysis of word+word compounds, according to which these constructions, which include more than one prosodic word, form a phonological phrase in a number of European languages.
72. Facing the problems noticed above, presented before in Vigário (1999a, 2000a), Galves (2000) has recently abandoned this idea.
73. See also Truckenbrodt (1999: 1.3) for a detailed argumentation against the visibility of empty syntactic elements and their projections in the mapping between syntax and phonology.
74. Three other studies that bear on issues related to the organization of the grammar in Portuguese are Lee (1995) and Schwindt (2000), on morphology and lexical phonology of the Brazilian variety, and Villalva (1994), on EP, but conducted from a morphological point of view.
75. As said in the preceding section, the fact that *-mente* and *z*-evaluative suffixes are independent prosodic words seems to be the crucial reason for these constructions to belong to the postlexical level (see Mateus and d'Andrade 2000: 97).
76. D'Andrade and Viana note in addition that the glide is obligatory with some numerals. In chapter 6, we show that this is a more general behavior as well, which, like the preceding cases, is also explained by the prosodic configuration that the relevant constructions display. Non-back vowel deletion is particularly difficult to grasp because it interacts with different types of phenomena, as we will see later in this book: the prosodization of clitics and of compound-like sequences (see chapter 5, section 4.3 and chapter 6, section 3); the reduction of clitic function words (see chapter 5, section 4.1.2 and chapter 7); and the lexicalization of combinations of function words (see chapter 7, section 5.4.2).
77. By adopting Vigário's (1998b, 1999a) analysis of non-back vowel deletion as a prosodic word limit process, Mateus and d'Andrade (2000) constitutes an exception to the generalized absence of the prosodic word domain from EP phonological descriptions. Frota (2000) includes this constituent in the formalization of a number of rules that apply within the intonational phrase domain. However, the prosodic word is not the topic under research. This contrasts with the other prosodic domains, which have been subject to some analysis. For example, (i) in d'Andrade and Viana (1994), Vigário and Falé (1994), Mateus (1995), Freitas (1997), Mateus and d'Andrade (1998, 2000), we can find descriptions of EP syllables, and in most cases proposals concerning syllable structure and the principles underlying its construction; (ii) d'Andrade and Viana (1989), Brandão de Carvalho (1989), Pereira (1990, 1999), and d'Andrade and Laks (1992) refer to the construction of (a constituent similar to) the foot; (iii) Frota (1996, 2000) studies the phonological

phrase and the intonational phrase, observing segmental, durational, intonational, and rhythmic phenomena; Falé (1995) and Vigário (1998a) also resort to the phonological phrase and to the intonational phrase in their account of intonational phenomena in this language.

78. We are using the traditional articulatory classification of EP segments (the terms palatal and velar are also used instead of front and back, respectively). In the standard variety of EP, only the back vowels [u, o, ɔ] are rounded. For a classification in terms of distinctive features, as defined in Chomsky and Halle (1968), see, for example, Mateus (1975: 72).
79. For extensive discussion and analysis on the data described here see, in particular, d'Andrade (1988, 1997), d'Andrade and Laks (1992), Pereira (1999) and Mateus and d'Andrade (2000: chap.6). In earlier accounts, stress is computed from right to left, constituents are binary, and left-headed, and the head of the rightmost constituent is assigned the primary stress. In all of these analyses, some vowels are exceptionally marked in a given lexical entry, so that they are not assigned stress. In Mateus and d'Andrade (2000) phonetically null class-markers are also assigned a rhythmic position. Within the verb system, most tense suffixes are usually claimed to be lexically marked in order not to bear word stress.
80. In the examples, whenever relevant we mark the vowel that bears word stress with a capital letter. For ease of exposition, we use in general EP orthographic system, and only when justified we employ the International Phonetic Alphabet, inside square brackets or slashes, to indicate phonetic or phonological forms, respectively.
81. There are also triplets of words that have the same underlying segmental structure and differ with respect to stress placement, as *sabiÁ* (a kind of bird), *sabIa* '(he) knew', *sÁbia* 'wise woman'. In this case, the words belong to different morphological classes: respectively, nominal, verbal, and adjectival classes. As pointed out by Mateus and d'Andrade (2000: 6.3), this sort of data argues against the view that EP stress is quantity sensitive (as claimed, for example, by Brandão de Carvalho 1989).
82. In this chapter we use the following abbreviations: PAST: indicative past perfect; IMP: indicative past imperfect; SUBJ: subjunctive present; 1/2/3Sg: 1st/2nd/3rd/person singular; 1/2/3Pl: 1st/2nd/3rd person plural; MASC: masculine; FEM: feminine; PL: plural; DIM: diminutive. Notice that in some of the examples in (66) the theme vowel does not surface due to the application of a rule specific of the verb system that deletes the theme vowel when it is followed by another vowel (see section 4).
83. We should add that enclitics behave like post-tonic syllables in that they are never assigned secondary stress in EP. This is already observed in Gonçalves Viana (1883), who notes the contrast between (some dialects of) Spanish and EP with respect to posttonic secondary stress assignment, as illustrated in (i), taken from Gonçalves Viana (1883: 62) (boldface signals secondary stress).

(i) PORTUGUESE	SPANISH	
<i>tŨmulo</i>	<i>tUmulo</i>	‘grave’
<i>dizIa-lhe</i>	<i>dicIale</i>	‘(he) used to tell him’

84. “Vowel reduction” is the traditional term used in EP literature for the phenomenon under observation in this section. We should notice that schwa, traditionally transcribed as [ə] in the literature on EP, is a very reduced central vowel. In the words of Gonçalves Viana (1883: 4): “[c’]est un *e* muet, comme on l’appelle généralement, bien plus étouffée, bien plus fermé, cependent, que l’*e* français de *me, le*.”
85. Brandão de Carvalho (1989, 1994) considers VR a regular phenomenon, which applies to *light* syllables. As for the vocalic nuclei in open syllables that fail to undergo VR, they are claimed to be *heavy*. The examples in (74) seem, nevertheless, to pose a problem for the assumption that VR is a general phenomenon, since the same underlying vowel, in the same syllabic environment, undergoes VR in some word-forms and surfaces unreduced in other (related) word-forms. In addition to this, under this approach the fact that syllables closed by /s/ or /r/ in non-final position do not prevent vowel reduction implies the postulation that such syllables are light, unlike those closed by /l/. Finally, there seems to be no other evidence for a distinction among vowels based on weight in EP. To be more accurate, word stress assignment is argued to be sensitive to weight by some authors (e.g. Brandão de Carvalho 1989, followed in Morales-Front and Holt 1997). However, Mateus and d’Andrade (2000: 6.3) and Pereira (1999) present a number of arguments against this view (see also section 2).
86. The rule that centralizes /i/ formalized in Mateus (1975: 26), d’Andrade (1977: 46), Andrade (1980: 48) includes a(n optional) final consonant. This formulation is meant to account for the final [i] in verb forms such as *pArtes* ‘(you) leave’, *Abres* ‘(you) open’. Again, this centralization seems to occur only in verbs, for in the noun system there are several instances of final unstressed [i] followed by a consonant (cf. i).

(i) <i>lÁpis</i>	‘pencil’	<i>Útil</i>	‘useful’	<i>mÁrtir</i>	‘martyr’
<i>oÁsis</i>	‘oasis’	<i>fÁcil</i>	‘easy’		
<i>grÁtis</i>	‘free’	<i>diffícil</i>	‘hard’		

It should be noticed that in early generative accounts some of the words presented in (i), namely those superficially ending in liquid, are assumed to have an underlying final /e/ (cf. d’Andrade 1977; Andrade 1980). Observing that the centralization of stressless final /i/ must follow the deletion of final stressless /e/, which is argued to apply very early in the derivation, Andrade (1980: 49) concludes that the centralization that affects /i/ must be specific to the verbal system.

87. It is often mentioned in the literature that words ending in stressless [i] are, nevertheless, rare in EP. We believe this is not because of vowel quality

facts, but rather results from the exceptional stress pattern that these words display. Since such words include no overt class marker, and given that stress in the nominal system usually falls on the last vowel of the stem (see section 2), these forms must be lexically marked in order for the stress to fall on the penultimate syllable, unlike in words such as *colibrí* 'id.', *saguí* 'tamarin', *javali* 'wild pig'. Finally, we should point out that there are also a number of stressed prefixes ending in a stressless /i/ where centralization does not apply (cf. chapter 6, section 3.2).

88. Diphthongs formed by a central vowel and [w] may constitute an exception to this generalization. For example, *saudar*, may be realized as [səwdár], although the most common realization of this word is [sodár]. As for the latter possibility, the most straightforward analysis is that an assimilation process has applied, changing [aw] to [ow] (the same type of assimilation is observable, for example, in the verbal system – see note 96). Thus, in the latter case VR does not apply because [o] is part of a diphthong. The realization of [ow] as [o], in turn, is a more general process of the Lisbon variety of EP (e.g. Mateus 1975: 42). The assimilation analysis may also explain the typical realizations of words like *saudade* 'nostalgia' and *saudoso* 'longing' with a [o].
89. The regular realization of these vowels as low is also referred in section 8.
90. The realization of initial vowels is discussed in section 10.
91. This description excludes stressless vowels followed by a fricative in word final position, since they surface as reduced vowels, as shown in (i).

(i) <i>ourIves</i> [i]/0	<i>ANtes</i> [i]/0	<i>alfEres</i> [i]/0
'goldsmith'	'before'	'second lieutenant'

We discuss the obligatory low realization of vowels in word final syllables closed by sonorants in section 8.

92. Only the compound linking vowel (–o) is usually reported to escape VR (e.g. Villalva 1994). The forms in (83) suggest, however, that this is a more general behavior of non-high vowels (–o or –e) in non-final stems of morphological compounds (see also chapter 6, section 3.3).
93. See chapter 6, sections 4 on clipping in EP. Notice that the adverb *inclusíve* 'comprehensively', which coexists with *inclusivamente* and is realized by many speakers with a low final vowel may also be analyzed as an instance of clipping. We should further point out that, to our knowledge, the generalization described in (vii) was not previously noticed in the literature.
94. Mateus (1997a) assigns vowel reduction to the postlexical component. The argument for that claim is based on the behavior of unstressed vowels in future and conditional tenses: when a clitic pronoun is inserted between two morpheme boundaries, stress is assigned as if the verb form contains two words and the stressed vowels do not undergo vowel reduction (e.g. *falAr-te-Ei*, [fəlártjɛj] '(I) will speak to you'); by contrast, when no clitic occurs, vowel reduction applies regularly to unstressed vowels (e.g. *falarEi*, [fələɾɛj]

‘(I) will speak’). This view presupposes that the verb forms inserted in the future and conditional are underlyingly identical whether the clitic is present or not. In chapter 4, we show that the two cases are obtained through different mechanisms, which explain the different behavior of the relevant vowels with respect to vowel reduction, while maintaining the assumption that this is a lexical phenomenon.

95. We assume the underlying form of the 1st person singular suffix to be /u/ rather than /o/ because this vowel always surfaces as high, and there is no evidence for its mid status (cf. also d’Andrade 1977: chap.4). The same applies for the underlying form of the nominal class marker –o (see Andrade 1980: note 37 for a similar remark).
96. In the case of 1st and 3rd person singular of the perfect indicative the theme vowel is not deleted possibly because it is followed by a glide: in the 1st conjugation, the theme vowel /a/ undergoes a specific process of assimilation under the influence of the glide, both in the 1st and in the 3rd person singular; in the 2nd and 3rd conjugations the theme vowel and the following glide undergo total assimilation in the 1st person singular, and thus a [i] surfaces; in the 3rd person singular, the theme vowel is not deleted also because it is followed by a glide. This is shown in (i) (for a similar analysis see e.g. Mateus 1983: note 18).

(i) <i>falei</i> < /fal a + j/	<i>bati</i> < /bat e + j/	<i>parti</i> < /part i + j/
‘(I) spoke’	‘(I) hit’	‘(I) left’
<i>falou</i> < /fal a + w/	<i>bateu</i> < /bat e + w/	<i>partiu</i> < /part i + w/
‘(he) spoke’	‘(he) hit’	‘(he) left’

We should notice, in addition, that fusion of a stressed high vowel and a palatal glide also occurs in non-verbal environments, as in *funis* ‘funnel’ (< funijs < /funil+s/) (cf. Morales-Front and Holt 1997), and the special case of assimilation involving the sequence /aj/ has also happened in other periods of the language in words such as *leite* (lat. *lacte*), and in the suffix –*eiro*, which is etymologically related to the suffix –*ário*. In Mateus and d’Andrade (2000: 4.2.2.1.2) a slightly different proposal is made, since these authors assume that EP has no underlying glides. Theme vowel deletion is claimed to apply only in the present tenses and thus the fact that the theme vowel is followed by a suffix starting with a vowel in the past tenses does not trigger vowel deletion.

97. The theme is defined as the verb stem plus the theme vowel.
98. We believe, however, that the nasalization of the vowel is possible in at least some of the cases where the prefix *en-* is followed by a nasal consonant: *en+madeirado* ‘in between wood’; *en+muralhado* ‘in between walls’. It is possible that speakers show some variability in these cases.
99. This analysis, proposed in Morales-Front and Holt (1997), suggests that the final fricative is syllabified in coda position in the lexical component. Thus,

the glide surfaces as non-back, even though the fricative may end up as the onset of the following syllable postlexically. The same is true for words ending in /l/ that undergo pluralization. Notice, nevertheless, that coda fricatives need not to be fully specified in the lexicon, since, if these consonants are simply specified as [-back], the right result with respect to the palatal glide is obtained. See Mateus and d'Andrade (2000: 2.4.2) on the underspecification of coda consonants.

100. In the case of words ending in stressed nasal vowel, as *fim* 'end' and *atum* 'tuna', we are assuming that the final glide, which has the same value for the features [back] and [high] as the preceding vowel, fuses with that vowel (the same type of process seems operative in the verbal system – see note 96). In forms like *patrão/patrões* 'boss(es)', also morphologically related to words like *patronal* 'relative to bosses', we assume that the underlying form of the stem is /patro[+nas]/ and the /o/ dissimilates to [ø] when followed by the nasal glide [w̃]. Notice that oral diphthongs of the form [ow] do not surface as such in the Lisbon variety of EP (e.g. Mateus 1975: 42) and there are no surface nasal diphthongs of this form either.
101. Some forms also exist where nasality is absent in etymologically related forms: e.g. *jejum/jejuar* 'fast/(to) fast', *dom/doar* 'gift/(to)give' (but *donativo* 'donation'), *ladrão/ladroagem* 'thief/thievery', *ordem/ordeiro* 'order/peaceful' (but *ordenar* '(to) order'), *tom/toar* 'tone/(to) sound' (but *tonal* 'id.'), *lua/lunar* 'moon/lunar', *um/uma* 'indefinite article-MASC/-FEM', *desdém/desdenhar* 'disdain/(to)disdain'. These forms are probably not to be synchronically related. The deletion of intervocalic –n– was a general phenomenon in Xth/XIth century but it was lost after the XIIth century (cf. Williams 1938; Teyssier 1980). Thus, these words can be assumed to have entered the language in different periods of time. See, nevertheless, d'Andrade (1977), d'Andrade and Kihm (1988) and Mateus and d'Andrade (2000: 7.2) for alternative views.
102. As said above, in Mateus (1975) and Mateus and d'Andrade (2000) the glide is inserted. We believe a simpler analysis is to allow for the underlying nasal segment to be specified as a non-syllabic nasal segment, which may surface as a nasal glide in word final position.
103. Notice that fusion must precede nasal vowel rise (cf. Mateus 1975: 1.4), for otherwise a low central vowel would surface, as it happens across words (see also section 13). This, again, suggests that the realization of the nasal segment is a lexical phenomenon.
104. The fact that there are some cases where the nasal segment surfaces as a consonant rather than a glide, as in (i), can be analyzed as involving an exceptional nasal consonant that is fully specified underlyingly.

(i) <i>abdómen</i> 'abdomen'	<i>cólon</i> 'colon'	<i>íman</i> 'magnet'
<i>dólmen</i> 'dolmen'	<i>cláxon</i> 'horn'	<i>slogan</i> 'slogan'

Under this view, these cases do not constitute an exception to final nasal diphthongization.

105. The same rule is also proposed in Andrade (1980: 57–58).
106. We assume that this process applies within a phonological domain, since it does not seem to refer to any morphological information.
107. Vowel harmony is a process specific for the verbal system that consists of the regressive long-distance assimilation of the height features of the theme vowel by the final vowel of the verb root. It only occurs when the theme vowel is deleted (e.g. Mateus 1975; d'Andrade 1994d; Mateus and d'Andrade 2000).
108. Although Mateus and d'Andrade (2000) do not analyze /e/-centralization in detail, they assign it to the postlexical component.
109. The syllabification of this glide contrasts with syllabification of the glide that is inserted to break a hiatus formed by oral vowels, as in *areia* 'send' (see section 7 below). In the latter case, according to Mateus and d'Andrade (2000: 57) the glide belongs to the rhyme of the syllable headed by /e/ (see also Gonçalves Viana 1883, for the syllabification of words like *maior* (mai.or) 'bigger'). Only in the latter case, therefore, the palatal glide provides the context for tautosyllabic /e/-centralization.
110. In the inflected forms inside brackets given in (111), the front vowel is stressless and there is no glide insertion. Instead, the front vowel followed by another vowel semivocalizes. This is also shown in (113) below.
111. As reported in section 6, an alternative pronunciation of the forms in (112) admits a glide. However, this glide behaves differently in some respects: (i) it is totally optional, (ii) it does not trigger /e/-centralization, (iii) it belongs to the onset of the syllable that follows /e/, rather than to the preceding coda.
112. We would like to thank an anonymous reviewer for bringing this point to our attention.
113. As noted in Andrade (1980: 4.2.1), this is true for non-round vowels. Round vowels show more variability, as illustrated in (i).

- | | | |
|------------------------------|-----------------|--------------------------------------|
| (i) <i>emoldurar</i> [u]/[o] | '(to) frame' | (<i>mOlde</i> [ɔ] 'mould') |
| <i>soltar</i> [o] | '(to) set free' | (<i>sOlta</i> [o] '(he) sets free') |

In order to account for the general realization of unround vowels in syllables closed by /l/ as low, Andrade proposes that the lowering rule also applies in these cases.

114. We should point out that schwas are often deleted in EP (see in particular chapter 7 of this book).
115. Notice that this may be conceived as a feature specification process, rather than a feature changing one, since non-high non-central vowels in this position may be underlyingly underspecified for the low feature. We should also add that there is no neutralization in the case of central vowels, since non-high central vowels are underlyingly low.

116. Notice, nevertheless, that d'Andrade (1977) assumes that words ending superficially in a liquid have an underlying final /e/.
117. To be more accurate, we believe that there are two possible syllabifications for these sequences: *pVrs* or *prVs*, where V stands for a V-position that may be filled by [i]. In the first case, the sequence [rʃ] belongs to the rhyme, whereas in the second, [r], like [p], belongs to the onset, and [ʃ] belongs to the rhyme (see also the remarks made in Morais Barbosa 1965: 5.16). What we believe is impossible is the separation of this sequence in two different syllables, as would be the case if an empty V-position existed between each of these consonants (*pV.rVs.). The latter syllabification seems to be allowed in Mateus and d'Andrade's proposal on syllable construction. External evidence for our claim comes from the difficulty speakers show in writing such words. While the alternation between the spellings *perspectiva* 'perspective' and *prespectiva* is frequent, the spelling *perespectiva* does not occur (unlike in other cases where a V-position may be posited, as in *segmento/segumento* 'segment', cf. Vigário and Falé 1994).
118. /r/ behaves differently from other segments that appear in coda (i.e. /l/, the nasal segment, and /s/): /r/ undergoes strengthening processes in word initial position and in the initial position of syllables preceded by a consonant (cf. section 9); /r/ does not semivocalize in word final position, as the nasal segment, or in word final position followed by a fricative, as the lateral and the nasal segment (cf. section 5); /r/ in word final stressed syllables may often be deleted (unlike lateral consonants) (cf. Leite de Vasconcellos 1901: 98). As there is an important generalization that there is only one consonant per rhyme in EP (e.g. Mateus and d'Andrade 2000), it follows that /r/ seems not to pattern like a 'full' consonant. This topic deserves further attention. See also Freitas (1995) for the problems involved in the acquisition of trills in EP.
119. Even if the second schwa is deleted, there is a clear perceptual effect concerning syllable structure: instead of the syllable division *li.dVrs*, if the vowel reduces speakers hear *li.drVs*. That is, in this case again the vowel does not behave as belonging to a syllable closed by /r/ in word final position.
120. For the sake of simplicity, from now on we will refer to the uvular consonant in its trill realization ([ʀ]), rather than in its fricative realization ([ʁ]).
121. Notice that the "onset rule" does not apply to all cases where /r/ appears in onset position. In the second position of a consonant cluster, as well as in intervocalic position, /r/ is subject to the default process, rather than to the "onset rule".
122. The absence of sonorant palatal consonants preceded by /r/ deserves further investigation.

370 *Notes*

123. Besides the impossibility of syllables closed by /r/, /l/ and /s/ in prefinal syllables of proparoxiton words, syllables with diphthongs and nasal vowels are also disallowed in post-tonic non-final position (see also Reighard and Almeida 1983, cited in Morales-Front and Holt 1997).
124. Since we found no cases of morphological conditioning of this process, we assume it applies at the left-edge of the prosodic word.
125. In a study on the acquisition of trills in EP, Freitas (1995) shows that the problems related to the production of these segments are not solved until the last stages of phonological development. In addition, Freitas reports cases where word initial [R] is substituted for [r] (e.g. *rei* 'king', *rato* 'mouse'). Although we believe that child production data do not necessarily mirror the adult system, and also despite the fact that in other positions the trill may be substituted for other segments (namely, [ʎ], [l] and [j]), these facts may lend further support for the derived status of word initial [R] in EP.
126. In the examples, we also show cases of English [r] in other positions of the word, where it is commonly identified with the EP alveolar tap.
127. These facts further suggest that the nasal segment is basically linked to the coda, before it is associated to the nucleus of its syllable. Thus, it is probably not basically a non-attached floating autosegment, as proposed in d'Andrade and Kihm (1988).
128. Glides in falling diphthongs are shown in d'Andrade and Viana (1994) to be part of the nucleus rather than of the coda position. An important argument in favor of this analysis is the fact that glides in falling diphthongs are nasalized along with the preceding vowel when followed by a nasal segment in the same syllable. The assumption that the glide is part of the nucleus and that the nasal segment spreads to this syllabic position, as proposed in d'Andrade and Kihm (1988), nicely accounts for these facts.
129. This issue is investigated from a sociolinguistic point of view in Mascarenhas (1997) (and more recently in Rodrigues 2001). Mascarenhas shows that several sociolinguistic factors correlate with the specific realizations of word initial vowels (e.g. 'sex', and 'region'). However, no integrated phonological analysis is reached. We will describe the process following our intuitions, which match in general the realizations of Mascarenhas' speakers that belong to the same sociolinguistic group as ourselves. In Mateus (1975: 4.2) and Andrade (1980: 4.5), different descriptions of initial vowel realization can be found. Although not all the facts described in this section are given in Mateus and d'Andrade (2000: 3.3.2), our description generally agrees with the one found in this work.
130. The description of initial vowels realization found in Gonçalves Viana (1883) points in this direction: front vowels are reported to always surface as high, while today they can also surface as mid; and initial back vowels in closed syllables are characterized as tending to be low, while this tends not to be the case in present day EP. In Leite de Vasconcellos (1901), in addition, round

vowels are reported to tend to surface as high, regardless of dialect considerations, while they usually surface as mid or low in present day EP spoken in the region of Lisbon. See also Morais Barbosa (1988) for an overview of different descriptions of the realization of initial vowels since the 18th century.

131. Forms like [ɛ]dito-[e]/[i]dital ‘edict/decree’ or [ɛ]timo-[e]/[i]timológico ‘etym/etymological’ seem to suggest that front low vowels may in some cases raise to [i]. However, in these forms a low stressed vowel does not necessarily imply that there is an *underlying* low vowel. In fact, as pointed out by Wetzels (1992) non-high stressed vowels of proparoxiton words are subject to a lowering rule and thus these vowels typically surface as low. This rule appears to apply without exceptions to stressed vowels in word initial positions of proparoxiton words. Therefore, both *édito* and *étimo* may be assumed to have an underlying /e/. Under this analysis, the alternative realizations in stressless position are the expected pattern if raising in one degree optionally applies to /e/.
132. The stressed mid vowel in *ovo* and *olho* is historically the result of metaphony, which may be described as a process of long distance assimilation whereby a low vowel in stressed position becomes mid if the following (adjacent) word final syllable ends in –o (for a detailed analysis, see d’Andrade 1994c, and references therein). The fact that the mid vowel in *ovito* and *olhito* occurs in non-stressed position and is followed by a non-final syllable, and the fact that the low realization is impossible suggests that the vowel has been reanalyzed as underlyingly mid. Notice, furthermore, that metaphony is no longer a productive rule. We therefore assume that the mid vowel is present in the underlying representation of these forms. Alternatively, we have to assume that the process that explains the realization of these vowels follows metaphony.
133. Given the lack of evidence that the initial position is defined morphologically, we assume that the relevant domain for the application of this rule is phonological (see also the discussion further below). We should further point out that the rule in (139) may, in fact, be seen as a trivial instance of Vowel Reduction. That is, unlike front and round vowels, central vowels pattern word initially in the same way as word internally.
134. We are assuming the feature geometry proposed in Mateus (1997b) and Mateus and d’Andrade (2000), developed within Clements (1988) and Clements and Hume (1995) framework. For the basic tenets of autosegmental phonology see, for example, Goldsmith (1991). We should stress that we are also assuming that the processes under observation apply to vowels in the initial position of a prosodic domain. This issue is discussed below.
135. In fact, the delinking of the [+low] feature may also affect the central vowel (and therefore no specification regarding the place feature is required in the formalization of the process). However, this vowel does not show the alternation between low and non-low realizations. Since they are obligatorily

- specified as [-low], in this case feature specification should apply in the lexical level, unlike in the case of non-central vowels.
136. The postlexical feature specification is possibly conditioned by sociolinguistic factors (see Mascarenhas 1998; Rodrigues 2001).
 137. In some dialects, the items in (150b) are usually produced with two vowels instead of a falling diphthong (see, for example, Gonçalves Viana 1883; Sá Nogueira 1938). This fact also suggests that the glides in (150) are obtained by a process distinct from the one involved in the forms in (149).
 138. D'Andrade (1998) claims that in a sequence VV where the first one is stressed and the second is high, the second vowel always surfaces as a glide. He notes furthermore that, in some dialects, the second vowel may surface as a vowel, but he claims that this is specific of the verbal system and is dependent on a tense morpheme boundary. However, the examples presented above involving a noun, and the forms involving a verb and a clitic (as in *vi-o* '(I) saw-him'), show that the second vowel may surface as a glide or as a vowel in prosodic word final position, regardless of morphological considerations (see also chapter 5, section 4.3.1).
 139. As mentioned in Câmara (1972), Gonçalves Viana (1883) presents (near) minimal pairs, such as *paulAda* ([aw]) 'blow' < *pAu* ([áw]) 'stick' and *paulAda* ([əu]) 'region of swamps' < *paUl* ([əú]) 'swamp', which are revealing in this regard.
 140. In fact, it is not impossible to conceive that the schwa surfaces as a glide, but in our dialect typically as a central but not a front one (a glide that exists, for instance, in Dutch – cf. L. Wetzels, personal communication). Nevertheless, in all the EP descriptions we are aware of, such a glide was never reported. We believe that the realization of the stressless final front vowel preceded by another vowel deserves further investigation.
 141. Similar facts are also reported by d'Andrade and Viana (1993: note 3).
 142. The rule is argued to apply within a *compound intonational phrase* (I^{max}). This domain is defined as a constituent that dominates two constituents of the same type (I) (see Frota 2000 for the arguments that support the existence of such a domain in EP, and chapter 1, section 5 of this book). We should notice that the upper domain that bounds this process is not crucial for our purposes here. The conditions for V1 semivocalization blocking, related to the presence of different levels of stress in the second vowel, are discussed in Frota (1996, 2000) and Ellison and Viana (1996). This issue is not relevant for us either and thus will not develop it here.
 143. See d'Andrade and Viana (1993), Frota (1996, 2000) and Ellison and Viana (1996) for further illustration.
 144. Notice that in cases such as *cuidadoso* ([uj]) 'careful', which has a falling diphthong, the second segment never occurs in a context where it surfaces as a vowel, and the glide is obligatory. In these cases, therefore, there is no reason to consider that the glide is obtained through V2 semivocalization.

145. On verbs ending in an underlying /i/, see section 3.
146. Although we have not investigated this issue in depth, our view of these facts is similar to the one presented for the raising processes that affect word initial stressless vowels. Specifically, we believe that non-high front stressless vowels (that is /e/ and /ɛ/) undergo a neutralization process that may be explained by the delinking of the features [low] and [high]. The specification of these lexically underspecified vowels as [+high] may be postlexical. As for centralization, the simplest analysis is to assume that it is also a postlexical phenomenon. This is so, because when these underlying vowels occur in stressless positions followed by other vowels, they usually surface as a non-back glide ([j]), as a consequence of V1 semivocalization (similarly to /i/). Thus, if they are lexically specified as non-back, the only feature that has to be modified is the syllabic one. This issue is left for future research. In the formalization of the rule below, we mark these segments as [-high] in order for the rule not to apply to segments that are lexically specified as [+high] (i.e. /i/).
147. To mention just two of the differences between Gonçalves Viana's (1883) data and most subsequent descriptions: Gonçalves Viana reports that a glide could surface if the following word starts with a stressless vowel, while this is not allowed in similar cases today; the presence/absence of the glide was dependent to some extent on V2 vowel quality, while today V2 quality does not seem to play any role in non-back non-high vowel deletion.
148. The judgments we provide in this section come from two sources: our own intuition and the observation of the production of four speakers in a reading task. The description of this task is presented in chapter 7. The most important difference between our description and some of the more recent observations on non-back non-high vowel deletion is that for us (and our informants), there is, in general, no glide whether the second vowel is stressed or not (as in 98b). Although judgements similar to ours are provided by Ellison and Viana (1996), the realization of the glide when V2 is stressed is given as the general case in d'Andrade and Viana (1993).
149. A different proposal is made in Ellison and Viana (1996) in a brief note on this issue. These authors suggest that when a glide surfaces it is the result of an insertion process. In Vigário (1998b: 4.1) we have extensively argued against this view, showing that the front vowel that may surface as a glide must be present underlyingly.
150. We should add that when the non-high front vowel is preceded by a vowel, deletion usually does not apply. Rather, a very short schwa-like segment is articulated (see note 140).
151. Notice that the possibility of clitics to surface without a glide will be argued in chapter 7 to follow from a more general property of clitics, namely the fact that they undergo phonological reduction.

152. On the prosodic conditions for V1 semivocalization blocking, see Ellison and Viana (1996) and Frota (1997, 2000). In these works, different analyses are also provided for the conditions on the blocking of round vowel deletion.
153. The data in Frota (2000), reproduced in (172), are based on actual productions of speakers in a reading task.
154. Notice that in cases such as (172d), Frota (2000) argues that the two I-phrases may be grouped together in a compound prosodic domain, and this is the domain within which the rule applies (on the notion of compound prosodic domain see also chapter 1, section 5 and note 157 below).
155. The minimal clash configuration is defined as a sequence of word stresses which may include one intervening unstressed syllable (Frota 2000: 91–92).
156. Vowel merger is described, among others, in Gonçalves Viana (1883), Sá Nogueira (1939), d' Andrade and Viana (1993), Ellison and Viana (1996), and Frota (1997, 2000). The upper domain that bounds this process is defined as the (compound) intonational phrase in Frota (2000), where its sensitivity to stress-clash configurations is also assessed. The span nature of the process is proposed in Vigário (1999a).
157. I^{\max} is defined as an intonational phrase that is dominated by a higher-level constituent. It may, however, dominate two constituents of the same level, in which case it forms a compound intonational phrase. This compound phrase is also reported to be the domain that bounds other rules. In addition, several facts show that the intonational phrases that are included in a compound phrase have also the properties of I-phrases: for instance, they also bound other phonological processes and set up the domain for tone assignment (see Frota 2000 for the complete argumentation and chapter 1, section 5 of this book for a review).
158. Notice that, in the cases illustrated in (178), Sá Nogueira (1938), like ourselves, does not consider V2 semivocalization possible either (see section 11.1). We should add that, in any event, these particular issues are not crucial for our analysis. In chapters 5 and 6, we will use this process as a diagnostic for the prosodic structure assigned to a number of constructions, considering only the segmental contexts where there is agreement in the literature on the possibility of central vowel deletion.
159. Some specific (very frequent) combinations admit syllable degemination with syllables headed by [v] (cf. *Avenida da...* 'avenue of...').
160. As it stands, we believe the conditions in (190) are capable of predicting the most clear cases of syllable degemination. Nevertheless, Mateus and d' Andrade (2000: 145) consider that in the sequence in (i), which violates the condition (iii), syllable degemination is possible.

(i) *disse sinceramente não* '(he) said sincerely no'

According to our intuitions, there is a strong tendency in these cases for the two consonants to surface as a long consonant, and thus what we have here is not syllable degemination but rather a sequence of two identical consonants

that become adjacent as a consequence of schwa deletion. The same tendency is also found when a round vowel is involved, as illustrated in (ii).

- (ii) *pouco capaz* ‘not very capable’
 posso saber ‘(I) may know’

We should further notice that Frota (2000) also reports the possibility of syllable degemination in a case with a stressed [o] in the second syllable (in the sequence *campo pouco* [pupó]/[pó] ‘field not-very’). Despite the variations in the judgments of these cases, what seems clear is that a non-high vowel is totally disallowed in the first syllable, whereas in the second syllable such a vowel is possibly tolerated.

161. Besides the conditions related to the phonological composition of the syllables involved in this process, Mateus and d’Andrade (2000) suggest that the presence of phrasal stress blocks syllable degemination. However, they do not elaborate on this issue.
162. In this preliminary study, Frota and Vigário (2000) analyze the prominence transcriptions of 80 sentences of EP, which were perceived by each of the authors (see chapter 5, section 4.3.2 for further details).
163. Although both d’Andrade and Viana (1989) and Frota and Vigário (2000) do not report any occurrence of the initial stress in the second syllable, this seem to be a possibility, although a much less common one, according to our informal observations of spontaneous speech. In d’Andrade and Laks (1992), d’Andrade and Viana (1989, 1999), and Pereira (1999) the alternation of stresses occurring in the first and the second syllable of a word is also described for words with an odd number of pretonic syllables. According to these authors, this follows from the interaction of two distinct mechanisms of non-primary stress assignment. The stress found in the first syllable is obtained by the positionally determined non-primary stress, while the stress on the second syllable is determined by the alternating rhythmic stress that is computed from the primary stress leftwards. The fact that all the rhythmic stresses between the second syllable and the primary stress are not perceived, however, suggests that the stress in the second position of the word is possible in some cases because it still counts as word initial. See chapter 5, section 4.3.2 for further discussion.
164. D’Andrade and Viana (1999: 91) claim that the initial position refers to the intonational unit (“unidade entoacional”) (on this notion, see chapter 2, section 4.2). They further argue that the domain of the initial stress rule is the intonational phrase rather than the word. In chapter 5, section 4.3.2 we challenge this view and propose that the domain of the initial stress rule is the prosodic word.
165. In the examples, we mark the syllables bearing emphatic stress with capital letters and underlining. Collected data on emphatic stress is presented in Appendix I.
166. In these examples, boldface indicates this specific type of stress.

167. Relying on our intuition and perception, we would characterize the initial emphatic stress as consisting of a local F0 rise on the initial syllable. In Vigário (1997a, 1998a) another type of stress is also reported to affect the negative adverb *não* 'not'. Since the emphatic stress falls on the initial position of words, and given that the negative adverb is monosyllabic, it is not clear whether the two phenomena are in fact the same. There is at least one difference that we can identify at this point between the two: in the case of the F0 rise in *não*, the F0 may remain high in the realization of subsequent pitch accents (cf. Vigário 1997a, 1998a), while we believe that in the case of initial prominence the F0 rise is necessarily a local phenomenon, affecting only the word initial syllable.
168. It is possible that there is a correlation between initial stress and emphatic stress. That is, it can be the case that emphatic stress is, in fact, a stronger (emphasized) realization of initial stress. Nevertheless, it is possible to have emphatic stress in a syllable that is immediately followed by the main stress of a word, which we believe is impossible for non-emphatic initial stress. This suggests these are two independent phenomena.
169. The pitch accents (that is, the tonal events that are associated with stressed positions) that are well established for EP are L*H, HL*, and H*L (see, in particular, Frota 1994b, 2000; Vigário 1998a). These phonological categories are related to the following phonetic properties: the vowel that is assigned an H is realized with a relatively high F0 value, whereas a vowel that is assigned an L is realized with a relatively low F0 value. The asterisk indicates that the preceding L or H is aligned with the stressed syllable. For the basic tenets of intonational theory this description is based on see, among others, Pierrehumbert and Beckman (1988) and Ladd (1996).
170. PA stands for pitch accent. Examples of the contours that instantiate the possibilities in (196bi) and (196bii) are, respectively, the sentences (77) and (76) of Vigário (1998a: 228).
171. The precise status of this initial H tone as a pitch accent or an edge tone is still an open issue. It seems to signal I-initial position, but, instead of being associated to the left edge of the I-constituent, it aligns with the right edge of the leftmost constituent of I. Putting aside the question of what the best analysis for this tone is, the initial H and its alignment with the post-tonic syllable of the sentence initial constituent are usually described in the literature.
172. The prosodization of clitics like *as* in (197) is discussed in chapter 5.
173. Earlier observations on the intonational properties of sentences with focused elements in EP can be found in Viana (1987: 4.3). Notice, nevertheless, that the notion of focus in Viana's work cannot be fully identified with the one presented in this section. For example, the contours (b–c) of the sentence "Quem deu um livro à Maria foi o Vasco" ('It was John who gave a book to Mary') in her Figure 2.18, are argued by the author to show three focused

- constituents in each case (cf. Viana 1987: 152), while, according to the characterization of focus made here, those contours could not be so analyzed.
174. On the notions of focus see, in particular, Jackendoff (1972), Rochemont and Culicover (1990), and the review in Frota (2000: 1.4).
 175. In the examples, the focused word is signaled with small caps.
 176. The phonetic details of the alignment of the H*L pitch accent, as opposed to the HL* pitch accent, are systematically investigated in Frota (2000: 5.3.2) and (2002b).
 177. The domain of range compression after the focus accent is suggested in Vigário (1998a: 148) to be the intonational phrase.
 178. See note 169 on some basic notions of intonational phonology used here. We should add that the subscript 'i' identifies a boundary tone, that is a tonal event that is aligned with the segmental material at the edges of prosodic constituents.
 179. In the case of heterosyllabic /e/-centralization and glide insertion to break a hiatus, no direct evidence was found for the process to apply with reference to the prosodic domain. In these cases, we have followed the intuitive idea that, in the unmarked cases, phonological rules apply within a domain defined by phonology, rather than by morphology. Recall also from sections 9 and 10 that /r/-strengthening and the particular realization of initial vowels are assumed to follow from strength properties of prosodic word initial position.
 180. As pointed out in chapter 2, section 4.2, Crysmann (1999, 2000a, 2000b) argues that bound pronouns in EP occupy a transitional place between morphology and syntax.
 181. The investigation in this chapter was partially presented in Vigário (1998c, 1999a, 1999c, 2000a).
 182. One additional property given in Zwicky and Pullum (1983) is not based on actual empirical observations, but rather follows from theoretical assumptions that we do not share: for Zwicky and Pullum, it is expected that syntactic rules may affect affixed words, but not clitic groups. We believe, in fact, that one of the properties that characterize a syntactic clitic is precisely its special syntactic distribution: on the one hand, they are manipulated by syntax in a fashion similar to independent words, and, on the other hand, they behave like bound elements that are syntactically dependent on another word. Thus, if a syntactic operation affects their host (like verb raising to Comp), it is expected that pronominal clitics are also raised with the verb, since they may cliticize to the verb in a lower position of the syntactic structure.
 183. In this chapter we use the following abbreviations. NP: nominal phrase; V: verb; N: noun; A: adjective; ADV: adverb; ART: definite article; PRES: indicative present; PAST: indicative past perfect; IMP: indicative past imperfect; SUBJ: subjunctive present; SUBJ IMP: subjunctive imperfect; FUT: future; COND: conditional; INF: infinitive; 1/2/3Sg: 1st/2nd/3rd/person singular; 1/2/3Pl: 1st/2nd/3rd person plural; ACC: accusative pronominal clitic; DAT:

dative pronominal clitic; IMPS: impersonal pronominal clitic; REF: reflexive pronominal clitic; MASC: masculine; FEM: feminine; PL: plural; PN: person/number suffix; TMA: tense/mood/aspect suffix; TV: theme vowel. In the pronominal system, gender is indicated only in the feminine forms.

184. For example, in Cameira (1994) the different accusative clitic forms are obtained from a single underlying form – /lu/. Consonant loss, in turn, is seen to follow from assimilation and deletion rules. In addition, it is suggested that these rules apply within the clitic group. In Barbosa (1996) this type of data is also assumed to be explained by the application of phonological rules, and is considered to constitute phonological evidence for the different directions of phonological cliticization of proclitics and enclitics (see chapter 2, section 4.2 for a review).
185. Although lexical affixes typically have fixed positions, i.e. they either attach to the left or to the right of their bases, there are at least two languages that have been claimed to have mobile affixes: Afar, a East Cushitic language, and Huave, a language of Oaxaca-Mexico (cf. Fulmer 1991 and Noyer 1994, respectively). Nevertheless, in none of these cases is affix location argued to be dependent on phrasal information, unlike in the case of EP pronominal clitics, as we will see in the next paragraph.
186. The relevant domain is tentatively defined in Frota and Vigário (1996) as the syntactic domain CP and the intonational phrase (I) prosodic domain. That is, the items that cause proclisis must be both within the same CP and within the same I as the clitic.
187. As a matter of fact, enclisis in contexts of proclisis is very frequent in younger generations, i.e. there is a recent regression of proclisis, as reported in Frota (1994) and Duarte, Matos, and Faria (1995). This shift in clitic placement follows naturally under the view presented therein that in this new grammar pronominal clitics are a step further towards the affixal status. Note that all the facts we are describing concern exclusively the (preceding) stage where proclisis and enclisis co-exist, and no claim is made here as to the reanalysis of pronominal clitics as inflectional affixes in future stages of EP.
188. We should add that, although interpolation is distributionally very limited (nowadays it is only found with negation, and possibly with a few other simplex adverbs), and it only occurs in certain styles and/or dialects, cases of interpolation can still arise in *lapsus linguae*, as documented in Frota (1994). This is suggestive of the psychological reality of a certain *mobility* of clitic pronouns in EP.
189. We thank Danièle Godard for drawing our attention to this point.
190. For the sake of comparison, all the examples show the same underlying vowel followed by the nasal segment (/e[+nas]/).
191. It should be noticed that there are a few cases involving compound structures which seem to have been reanalyzed as simple words so that final diphthongization resulted in simple nasalization of the preceding nucleus and deletion of the nasal segment, as in word internal position:

- (i) *sensaboria* [č]/??[č] (cf. *sem sabor* [č]) ‘without flavor’
 ‘tediousness’
 sempre-em-pé [č]/??[č] (cf. *em pé* [č]) ‘on foot’
 lit. always on foot

This is, in fact, what we would expect in verb final position if pronominal clitics had been reinterpreted as internal to the verb form.

192. Note that mesoclisism is not a counter-argument to this generalization, under the view presented in section 4.3.
193. This argument builds on a similar observation on the distinction between inflection and derivation, made in Anderson (1992: 79). On the notion of *portmanteau* see also, for example, Bauer (1987: 18).
194. As pointed out in Pesetsky (1985), Sproat (1988), and Booij and Lieber (1993), among many others, the existence of forms like *unhappier* give rise to a paradox: while from a phonological point of view, the word must be derived from *happier* (since *unhappy* would constitute a trisyllabic base), from a morphological point of view it must be derived from *unhappy*, given its semantic interpretation. See chapter 1, section 3.2 for a discussion on the resolution of such paradoxes.
195. It should be noticed, nevertheless, that there are many disyllabic bases to which *-ez* is added. Evaluative suffixes seem to exhibit a similar behavior, according to Skorge (1956), cited in Villalva (1994: 5.3.3): an evaluative form like *-inho* or *-ito* tends to select disyllabic bases, while *z*-evaluative suffixes like *-zinho* or *-zito* tend to select polysyllabic bases. As Villalva observes, there is, nevertheless, great variation in this case, as other factors also seem to play a role in the presence of one type of suffix or the other.
196. Crysmann (2000b) suggests that the fact that bound pronominals do not attach to participles is an instance of an “arbitrary gap”. We believe, on the contrary, that this *gap* may result from the properties of the participle itself instead of the clitic’s, as it is well-known that participles do not behave like full verbs in many respects.
197. In other syntactic accounts of cliticization, such as Sportiche (1992), pronominal cliticization is treated as inflection (see, for example, Manzini 1998 for a review of syntactic approaches to cliticization). Spencer (1991: 391–392) observes that syntactic treatments where *clitics* and inflection are conceived as abstract features associated to functional categories imply that it is necessary to find other, non-syntactic, criteria for distinguishing “full words from clitics from affixes from other morphological phenomena”.
198. Besides the syntactic information that plays a role in pronominal clitics distribution, EP clitic positioning is shown in Frota and Vigário (1996) to be sensitive to prosodic information. They suggest that proclisis is triggered by heavy function words that have to precede the clitic pronoun within the same CP and the same intonational phrase. Following this hypothesis, Duarte and Matos (2000) conclude that proclisis in EP is in general an instance of “*Move*

occurring between Spell-Out and the P-A Interface”. Note that, independent of the exact point in the grammar where clitic position is obtained, it still has to precede lexical insertion, as the form of clitics varies according to their position with respect to the verb (e.g. *come-lo* ‘(you) eat it’ vs. *não o comes* ‘(you) don’t eat it’).

199. Each of these forms may also appear in the feminine (*a/la/na*) and/or in the plural (*os, as/los,las/nos,nas*). For the sake of simplicity, in general we omit them, since they behave (*mutatis mutandis*) like the masculine singular form.
200. In grammars no distinction is usually made between nasal diphthongs found in the 3rd person plural and nasal diphthongs found in the 3rd person singular: in both cases the form of the clitic is said to be *no* (e.g. Cunha and Cintra 1984: 280).
201. There is another instance of consonant deletion, affecting words formed with the evaluative suffix *-zinho*: when the base ends in fricative, this consonant is obligatorily deleted. Thus, the diminutive form of *azuis* ‘blue-PI’ is *azuizinhos* and not *azuizinhos*. The first of two fricatives also tends to be deleted within the word (e.g. *transgênico* ‘transgenic’). Across words, on the contrary, the two fricatives may occur, as in *as zebras* ‘the zebras’, where [ʒz] or [ʒ] are possible realizations, but not [z]. We can add that the latter possibility, but not the penultimate one, may be found in more conservative dialects.
202. Notice that the lexical processes involving clitics and their hosts in Polish, referred above in section 3.2, can also be an instance of precompiled phrasal phonology (cf. Halpern 1991).
203. Cases of allomorphic alternations that are accounted for in a similar fashion are the alternation in the form of the English indefinite article *a/an* and in the form of Spanish feminine article *la/el* (cf. Hayes 1990).
204. For those speakers that do not differentiate between verb forms ending in nasal diphthongs, the relevant context for the insertion of the allomorph *no* should be [... [+nas]]_{vb} _ , instead. Actually, there are other instances of variation related to the selection of pronominal clitic allomorphs in EP. To mention just two examples, in some dialects, the allomorph *lo* is also selected in nasal contexts; *o*, instead of *lo*, may also be selected when the verb ends with a fricative, and in some cases, a linking [j] may appear between the clitic and a preceding fricative (cf. Mota 2000).
205. Notice that the relevant information assessed by these rules is exclusively morphosyntactic and segmental, not prosodic. Therefore, the application of these processes cannot support any claim related to the direction of phonological cliticization of pronominal clitics, as suggested in Barbosa (1996) (see chapter 2, section 4.2). In fact, nothing prevents the occurrence of precompiled phrasal rules between two elements that are not under the same prosodic node.

206. Certain combinations of prepositions and articles in EP, like in Italian (cf. Napoli and Nevis 1987) or French (cf. Zwicky 1987b), seem to be analyzable in a parallel way.
207. In the Perfect tense, TMA and PN are portmanteau forms (e.g. Villalva 1994) and therefore these are not considered in (236a).
208. The existence of two forms that alternate (a/e) fits well with the hypothesis that this element has a stem status. Indeed, while there are no cases of formal alternations involving inflectional affixes, there are several instances of stem alternations in EP within a single conjugation (e.g. *valho/vales* '(I/you) worth'; *trago/trazes* '(I/you) bring'; *faço/fazes* '(I/you) do' – e.g. Mateus 1975: 2.5).
209. Notice that the same rule could explain some specific phonological forms associated with dative and accusative clitic combinations: *mo(s)/ma(s)* (<me+o(s); me+a(s)); *to(s)/ta(s)* (<te+o(s); te+a(s)); *lho(s)/lha(s)* (<lhe+o(s); lhe+a(s)). Thus, clitic clusters, but not the sequence clitic plus *-haver*, show the behavior expected from inflected pronouns. Notice further that, according to some analysis, vowel deletion only applies in the present tenses, since in the past tenses, the absence of vowel is seen to follow from a rule of fusion (cf. chapter 3, section 4). In any event, we can establish the generalization that whenever two adjacent vowels appear in verbal inflectional environments, some process applies that prevents the adjacent vowels to surface. Therefore, also under this formulation of the facts, clitics do not behave like inflectional affixes.
210. We use the term *syntactic clitic* to express the fact that *-haver* has a syntactic distribution different from other *free* auxiliaries, as it is *dependent* on the infinitive verb+clitic sequence.
211. We follow the view presented in chapter 1, section 3.2, and argued in Booij (1988), Booij and Lieber (1993), and Nespor (1990), that prosodic structure up to the prosodic word level is built in the lexicon. See, in addition, chapter 5, section 4.2 for the arguments in favor of the lexical prosodization of clitics, and chapter 3, sections 2 and 3 respectively for the lexical status of stress assignment and vowel reduction in EP.
212. Reanalysis has presumably given raise to a portmanteau suffix. That is, the adjacent morphemes that realize the infinitive and the present/imperfect features were probably reanalyzed as a single unit. Thus, the two morphosyntactic sets of features are realized in a single suffix. In fact, in EP there are no instances of two co-existing TMA morphemes, thus suggesting that in this language there is only one slot for TMA markers in the inflectional template. Notice that, as we will see below, reanalysis is only possible in the synthetic form (the one without an intervening clitic), since in the analytic form, the two sets of morphosyntactic features are realized by two distinct formal units. See Villalva (1994: 4.3.3.2) for a proposal on the morphosyntactic features associated with tense, mood, and aspect in EP, and for a classification of portmanteau suffixes in the verbal system. Notice that the formation of future

- and conditional is excluded from Villalva's observations: because of the possible occurrence of internal clitics, these are viewed as compound forms (cf. Villalva 1994: chap.4, note 57).
213. For a similar remark on the cues for the prosodic word in English, see Raffelsiefen (1999a). Notice that, as mentioned in chapter 1, section 6.1, (re)syllabification is also frequently cited as a useful test for this prosodic domain in Germanic (but not in Romance) languages (e.g. Booij 1995, 1996a; Wiese 1996; Hall 1999a; Raffelsiefen 1999a).
 214. Notice, in addition, that word stress is assigned with reference to a domain morphologically defined, rather than with reference to the prosodic word domain (see chapter 3, section 2).
 215. This sort of processes may in fact exist in languages that impose minimality restrictions on prosodic words. An example of a rule argued to be motivated by minimality requirements is epenthesis in imperative forms of monosyllabic stems in Shona and KiNande (two bantu languages) – cf. Downing (1999) and references therein. Leben and Ahoua (1997: 126) also suggest that the inclusion of monosyllabic words into adjacent prosodic words in Baule is motivated by minimal word requirements.
 216. Recall that under the analysis presented in chapter 3, section 5 final nasal diphthongization consists of the semivocalization of an underspecified nasal segment in coda position, rather than on glide insertion.
 217. For this computation we have considered nouns, verbs, adjectives, adverbs, and stressed pronouns, and we have excluded function words – e.g. determiners, prepositions, pronominal clitics, conjunctions, complementizers and other particles – since some function words do not form independent prosodic words, as we will see in section 4.1.2.
 218. Note that the two corpora and methodology are not entirely comparable: we have excluded from the observation function words, which are often monosyllabic, and the words considered are inflected; whereas Thornton (1996) evaluates the proportion of monosyllabic nouns with respect to the whole set of words contained in the Italian Basic Vocabulary, regardless of their category.
 219. There are two possible realizations of this word, depending on the application of the semivocalization of the vowel preceding the schwa (see chapter 3, section 11.2 for the description of this optional process). Naturally, only in the realization where V1 does not semivocalize is schwa in syllable initial position.
 220. See Peperkamp (1997a) on the prosodic word readjustment caused by resyllabification (see also chapter 1, section 7 of this book for a review).
 221. Although it is generally accepted that there are two sources for schwa in EP (one resulting from the application of vowel reduction to non-back vowels and the other resulting from an insertion process), they are usually not explicitly differentiated. Nevertheless, the presence of schwa has been related to factors like speech style and/or rate (e.g. Mateus et al. 1990: 303).

222. In contrast with other Germanic languages, English allows for words to begin with a schwa (cf. Raffelsiefen 1999a: 148, among others).
223. See Nespor (1985) for a similar observation in Italian. It is also well-known that [ð] in English is only possible word initially with function words (e.g. Morgan, Shi, and Allopenna 1996, among many others).
224. We omit here pitch accent and focus assignment, since it may be thought that these phonological events are associated with prosodic word stressed syllables and thus may constitute primary evidence to word stress, rather than to the prosodic word, directly. See, however, chapter 6, in particular section 3.1, for the suggestion that pitch accent distribution as well as focal stress are sensitive to prominence relations defined within compound prosodic word domains. As for the I-initial H-tone, it usually aligns with the edge of the prosodic word domain (see chapter 3, section 18) and thus we will treat it as a boundary tone, which may therefore cue the prosodic word domain independent of word stress.
225. As we will see in chapter 6, section 3.1, in addition to the suffixes identified at least ever since Gonçalves Viana (1883) to bear word stress independently of their base (z-evaluative suffixes and the adverbial suffix *–mente*), we could find another affix-like unit with the same properties, which is *–avos*.
226. We assume that [r] is part of the inflectional affixes that mark the future and conditional (see the discussion in chapter 4, section 4.3). In the examples we use capital letters in underlying to signal the syllable that bears emphatic stress.
227. In EP, a schwa between consonants is frequently not realized (e.g. Morais Barbosa 1965; Mateus and Delgado Martins 1982), unless it receives initial or emphatic stress. Since when a schwa in this context is possible its deletion is usually also possible, for the sake of simplicity we will not mark the latter realization in the exemplification.
228. Notice that the realization of initial vowels remains variable despite the occurrence of resyllabification, which causes the prefix final consonant to become the onset of the syllable starting the morphological base. Resyllabification in EP, like in other Romance languages, has a phrasal domain (the intonational phrase) (cf. Vigário 1999a, and Frota 1998 for relevant acoustic data concerning resyllabification of final fricatives in EP).
229. The form *reiterar* ‘to restate; to emphasize’ is historically prefixed and the prefix is realized with a [ɐ]. We assume that a reanalysis of this word as a simple word has taken place, since *iterar* is not used as an independent word and consequently the meaning of the word is not compositional.
230. Notice that, according to Chomsky (1986: 7), in a structure where a constituent α is adjoined to a constituent β , β consists of two ‘segments’, and therefore in such configurations β does not dominate α .
231. Note that in the 118 cases of emphatic stress collected, only 3 show stress in an underlying non-initial vowel of a non-prefixed word. In 2 of these, in

- addition, the ‘stressed’ vowel corresponds to the first *phonetic* vowel of the relevant words (see note 262 below for further details).
232. For example, syllable degemination may apply in the sequence (*indique*)_ω *que terminou o serviço* ‘indicate that you have finished the work’. According to the formulation in (260), the process may apply regardless of the prosodization of the proclitic *que*, which, in theory, may attach to the following prosodic word, or to higher-level prosodic constituents (see chapter 1, section 6.3 on the different possibilities of prosodization of clitics). We owe this observation to Lisa Selkirk. We will see in section 4.3, nevertheless, that in EP all proclitics are adjoined to the following prosodic word.
 233. By contrast, emphatic stress seems to be optionally assigned both to the first and to the second vowel of the word *desodorizante*. Like in chapter 3, in this chapter we use the symbol ^x to indicate a very marginal/almost impossible realization.
 234. But see chapter 4 for the discussion on the affix/clitic status of pronominal clitics in EP, and the conclusion that they are not affixes.
 235. Notice, however, that in the case of simple clitics the presence of postlexical stress assignment blocks the occurrence of the clitic (reduced) form, whereas in the case of special clitics and bound words such stress assignment does not prevent the occurrence of the clitic words. Thus, the latter class of clitics may bear (postlexical) stress. Languages reported to have lexically unstressed clitics that may end up with a postlexically assigned stress include Greek (cf. Nespor and Vogel 1986; Berendsen 1986: 4.2; Nespor 1999b), French (cf. Van der Leeuw 1997; Nespor 1999b), Neapolitan (cf. Bafile 1994; Peperkamp 1997a; Nespor 1999b), and Southern Calabrian (cf. Nespor 1999b), among others.
 236. The expression ‘at most’ is intended to allow for clitics to be smaller than a syllable, as in the case of some clitics in English (cf. Selkirk 1984), Dutch (cf. Berendsen 1986; Booij 1996), or German (cf. Wiese 1996; Hall 1999a).
 237. For example, while auxiliary verbs and modals may be clitic in languages like English (e.g. Selkirk 1984, 1996), the elements that belong to these classes in Igbo bear word stress (cf. Zsiga 1992: 128). Similarly, some adverbials are clitic in Dutch (e.g. Booij 1996), but not in EP. Conversely, the definite article is a clitic in EP, whereas determiners in Korean form autonomous prosodic words (e.g. Jun 1996). On the identification of EP clitic words see section 4.1.2.
 238. To be more accurate, [v̥] may appear in stressed position but only when followed by a palatal segment (deriving from an underlying /e/) – see chapter 3, section 6 – or followed by an underlying nasal segment (deriving, in this case from an underlying /a/) – cf. Mateus (1975), among others.
 239. We assume that the obligatorily contracted forms involving prepositions and articles, as well as personal pronouns, are lexically listed as single units (see, chapter 4, section 4.2.2 for the analysis of EP personal pronouns, and, in

particular, Napoli and Nevis 1987 for the defense of a similar approach to certain preposition plus article clusters in Italian).

240. Clitic words ending in schwa may also undergo further reduction, that is, their only vowel may be deleted (see chapter 7). In addition, with enclitic schwa-words (i.e. postverbal pronouns) vowel deletion is obligatory. This is so because, as argued in section 4.3.1 below, enclitics are incorporated into the host's prosodic word and thus become prosodic word final. The incorporated clitics therefore create the environment for the application of the rule that deletes non-back vowels in prosodic word final position.
241. To be more accurate, we will see in section 4.3.1 that some function words can receive a pitch accent if they are assigned intonational phrase prominence. This seems only to occur with complementizers and relative pronouns in intonational phrase final position (see also Frota and Vigário 1996).
242. To be rigorous, the stressless form *que* may appear either preceded by *o* or not, while the stressed form *quê* must always be preceded by *o* (see Âmbar 1992 for a syntactic account of Wh-movement in EP and of the form of Wh-words in different contexts).
243. An alternative analysis would be to consider that all interrogative words have a strong and a weak form, although only in some cases we may distinguish these forms on the basis of their phonological make-up. There is, however, an important argument against this hypothesis: the behavior of *porque* and *onde* is distinct with respect to the process of final non-back vowel deletion. In fact, as shown in (i) below, this process usually applies to *onde* but not to *porque*. Notice that in the latter case the vowel may also be deleted, as in other cases of reduced clitic words, but deletion is not obligatory, as it normally happens with prosodic words (see chapter 7 on the distinction between the two processes and on collected data corroborating these observations).

- (i) a. *Porque abandonaste a sala?* [j]/0
 'Why did (you) abandon the room?'
 b. *Onde abandonaste o gato?* ^x[j]/0
 'Where did (you) abandon the cat?'

This shows that *onde* patterns like a prosodic word, unlike *porque*.

244. We have omitted the cases where a glide precedes the syllable nucleus, because in such environment glides are obtained through the application of V1 semivocalization, which is a postlexical process (cf. chapter 3, section 11.2).
245. We should add that EP syllables in full words may also include two consonants in the Rhyme, as in the word *pers.pi.caz* 'acute', although this is extremely rare (see also chapter 3, section 8).
246. In fact, some of these syllable structures may arise in reduced forms of clitics or in contractions of clitic clusters, as in (i).

386 Notes

- | | | | |
|---------|-----------|------|----------------|
| (i) CCV | (para) | pra | ‘for’ |
| CCVC | (para as) | prás | for-the-FEM-PL |

247. In section 4.3.2 below we justify why we do not include final round vowel deletion and final central vowel deletion in the list of processes that can be used to assess the left-edge of the prosodic word (see section 3.2 above on a similar discussion concerning syllable degemination).
248. The judgments of the forms in (285c) are difficult to establish because we have a sequence of a stressed vowel and a high unstressed vowel, and these stressless vowels may be extremely reduced – this is why we have used the symbol ^x to indicate that this realization would be nearly impossible. Given the difference in sonority between the two vowels, for some speakers the second vowel may be felt to become a glide. According to our judgments, however, it is not possible to syllabify the sentences in (285c) as *[éw.víw.tɛj̥f.tu], *[sɔ̃.v.tíw.díj̥f] but only as [éw.ví.u.tɛj̥f.tu], [sɔ̃.v.tí.u.díj̥f] (where the dots signal syllable boundaries). By contrast, both possibilities are available in the examples in (285b): [...ví.u...]/[...víw...] and [...lí.uʃ...]/[...líwʃ...].
249. The segmental composition of vowel initial clitics does not allow to test the application of the rule to the final vowel of a verb followed by a postverbal clitic. In fact, vowel initial pronominal clitics have either a central vowel or a high vowel. In such a configuration, as reported in chapter 3, section 14, the rule does not apply, independently of the prosodization of clitics. Despite this, it is possible to establish a contrast between the deletion of the central vowel when followed by a high vowel that belongs to a(n en)clitic, which is totally impossible (cf. ia), and the deletion of this vowel when followed by a high vowel that belongs to another full word, since the latter yield, nevertheless, a better result (cf. ib).
- | | | |
|----------------------------|---------------------|----------------------------------|
| (i) a. (isso,) diga-o você | [ɐ]/*0 | ‘(as for that,) say it yourself’ |
| b. casa usada | [ɐ]/ ^x 0 | ‘used house’ |
250. Another fact that suggests that this is not a pure phonological process is that in a few specific sequences involving very frequent function words, and in certain familiar registers, a sequence of vowels may also undergo merger to [ɔ] even though the first one is not [ɐ] or [a]. This is documented in (i), where some of the possible realizations are given.
- | | | |
|----------------------------|-----------------------|-------------------------|
| (i) todo o lado | [tódu]/[tóɔ] | ‘everywhere’ |
| como o João | [kómu]/[kómɔ] | ‘like (the) John’ |
| mais do que o Pedro | [dukju]/[duku]/[dukɔ] | ‘more than (the) Peter’ |

This sort of merger is in general impossible with other words (cf. ii).

(ii) eu <i>mando o</i> correio	[mẽ́duu]/*[mẽ́dɔ]	‘I send the mail’
eu <i>como o</i> gelado	[kómu]/*[kómɔ]	‘I eat the ice-cream’
<i>indique o</i> caminho	[ĩdíku]/*[ĩdíkɔ]	‘show the way’

The phenomenon is possibly conditioned by word frequency (see chapter 7 for other phenomena conditioned by word frequency).

251. In the case of *ao* we will consider the monophthongized form to be the result of a reduction process in chapter 7. It is possible that the same line of explanation may be pursued in other cases where clitics and perhaps other frequent words are involved.
252. As we have seen in section 4.1, the function word *pelo* may have reduced forms. In order to assess the application of syllable degemination it is important to select the unreduced form [pelu], since the reduced form [plu] shows a complex onset, and therefore does not provide the segmental context for the rule to apply (see chapter 3, section 15 for the details concerning the segmental conditions on this process).
253. Given the existence of solid arguments for the incorporation of postverbal pronominal clitics into the verb’s prosodic word, and since enclitics do not affect word stress location, we must conclude that word stress location is not a valid argument for determining the prosodization of clitics in EP. By assuming that word stress is a lexical rule (see chapter 3, section 2), we have actually predicted that the postlexical combination of clitics with their hosts could not affect word stress. In chapter 8, we review similar data of Standard Italian in the light of EP facts and discuss the consequence of our findings for the prosodization of enclitics in that language.
254. In fact, no sentence in EP may end in any of the function words considered in this chapter (to the exception of the form *cada* ‘each’ when it is used as a pronoun). In general, only complementizers may end an intonational phrase. Only when a parenthetical expression is inserted after a preposition (which is always rather marginal), or after a relative pronoun may such a function word occur in this configuration. In the case of the definite article this seems totally impossible. The reasons for this are possibly syntactic rather than prosodic (since, for example, demonstrative items like *este* ‘this’ are stressed, but may not be separated from their noun phrase either). These peculiarities fall outside the scope of the present work, and thus we will not develop them further.
255. This assertion is based on informal observations of uttered sentences by EP speakers, as well as on introspection. Although we believe this description does not raise any controversy, to our knowledge this question was never raised in EP literature and thus we found no reference to this distinction between complementizers and postverbal weak pronouns.

256. According to Bafile (1994), in Neapolitan verb-clitic sequences with two stresses, the first is a secondary stress, while the stress on the clitic is the primary one.
257. We will not explore here other possibilities for the phrasing of these sentences. It is however important to observe that alternative (preferred) prosodizations consist of the integration of the complementizers either into the following or into the preceding I-phrase, as already described in Frota and Vigário (1996) and Vigário and Frota (1998). The existence of the three possibilities may be due to the interaction between the mapping algorithms of I-construction, on the one hand, and the prosodic weight requirements on I-phrases, on the other hand (on these issues, see Frota 2000 and chapter 1, section 2 of this book for a review).
258. Notice that even under this postlexically assigned stress, the final vowel of *que* and *porque* surfaces as a schwa. In fact, these complementizers have no strong variant, unlike the homophone interrogative pronouns (see section 4.1.2).
259. The present discussion excludes the cases where a stressless word follows a stressed function word (like the emphatic expression *é que*, and *se bem que/já que/só que* ‘although’, among others). In fact, in these cases, it is possible that the cluster of function elements, which appear together very frequently, may have lexicalized as a single unit. These cases deserve a specific study, and thus are left for future investigation.
260. This behavior is reminiscent of the behavior of the Šrem dialect of Serbo-Croatian described in Selkirk (1996), where proclitics are assigned the same representation as the one we are proposing for EP (see chapter 1, section 6.3).
261. Gonçalves Viana (1883: 62) also rules-out the possibility of a *secondary stress* assigned to postverbal personal pronouns in EP.
262. Of the 118 attested tokens with emphatic stress in Appendix I, in 110 emphatic stress occurs in the first syllable of either a full word or of a clitic; in 5 cases it occurs in the second syllable of a prefixed word, which also counts as prosodic word initial (see section 3.2); in only three cases is the emphatic stress realized on a non-initial syllable. Notice that, even in the latter cases, it occurs in the first phonetic vowel in two tokens, and on the second phonetic vowel of the word in one token. These cases do not therefore constitute important counter-evidence to our proposals, which account straightforwardly for 97.5% of the collected data.
263. As noticed in chapter 3, section 6, heterosyllabic /e/-centralization does not operate when this vowel is in a stressless position and is followed by a palatal consonant. We have assumed that this is so because vowel reduction (i.e. raising and centralization) has operated. When the vowel is followed by a high palatal vowel, by contrast, vowel reduction is regularly blocked, and therefore the conditions for centralization to apply are met. There are no suffixes that start with a palatal consonant and are preceded by a stressed /e/, and thus the forms in (327) are not maximally comparable. We recall,

nevertheless, that in simple words, heterosyllabic /e/-centralization regularly applies in forms that are segmentally nearly identical to those involving clitics in (327b) above (cf. *tenho* [tẽpu] ‘I have’, *telhas* [télʁɐ] ‘tiles’ – see also chapter 4, section 3.2).

264. To be rigorous, there is one word in the list cited above that is historically derived with the prefix *re-* which is *reunião* [ʁjunjẽw] ‘meeting’ (see also section 3.2 on the phonology and semantic interpretation of this word form). Interestingly, as we will see in chapter 7, this particular word is sometimes realized with a very reduced glide, which, in extreme cases, may even be absent. This supports the present approach that glide deletion is correlated with the frequency of words rather than with the lexical/postlexical locus of prosodization of prefixes and proclitics.
265. We thank Geert Booij for drawing our attention to this point.
266. By contrast, prefixes may undergo postlexical span rules that have a domain higher than the prosodic word (e.g. V1 semivocalization), as well as postlexical phenomena that refer to the left-edge of the prosodic word domain (e.g. initial stress assignment) (see chapter 1, section 2, note 13 on the notions of span, limit and juncture rules). However, as expected, such phenomena also affect proclitics.
267. The expression “if any” in (333) is intended to cover the cases where there is no following prosodic word. This happens with clitics at the lexical component, since they are only combined with their hosts postlexically, as well as with proclitics at the postlexical component when they appear in intonational phrase final position (see section 4.3.1).
268. From the point of view of the morphological structure, according to Villalva’s (1994) approach, *-ção* is added to the verbal theme, yielding a nominal *theme*, headed by a theme index; to the theme the morphological inflection (phonetically empty in this case) is then attached, which finally creates the *word* of the nominal category.
269. Other languages where a stem plus any suffix are grouped together under the same prosodic word include Italian (Nespor and Vogel 1986; Peperkamp 1997a), French (Hannahs 1995b), and Spanish (Peperkamp 1997a).
270. See, in particular, Selkirk (1986, 1996) for an edge-based approach to prosodic domains construction.
271. The generalization that *Lex^o* is mapped onto a prosodic word will also play a major role in the prosodization of compound-like constructions (see chapter 6, section 5). A similar proposal has been put forward in Nespor and Ralli (1996) and Nespor (1999b) to account for the prosodization of compounds in several languages, although the consequences extracted from it are not precisely the same as ours (see chapter 6, section 5 and, for a review of this proposal, chapter 1, section 6.4).
272. Notice that, under the present view, pronominal proclitics end up adjoined to the verb’s prosodic word either as a consequence of the general condition of

rightwards adjunction, or as a consequence of the definition of the postlexical prosodic word: (i) if they are not part of the same Lex^o as the verbal host, they are prosodized like other clitic function words; (ii) if the proclitic is part of the syntactic head that includes the verb instead, the prosodic word boundary to the left of the verb is obtained because it comes from the lexical level (and left boundaries are postlexically projected), and the prosodic word boundary to the left of the clitic follows from the mapping of Lex^o nodes onto prosodic words.

273. Before proceeding, we should point out that it is not always possible to determine the behavior of a given construction with respect to all relevant phenomena. In fact, in some cases, no examples can be found where the context for a given rule either to apply or to be blocked occurs.
274. On the morphological status of *z*-evaluative suffixes and *–mente*, see, for example, Villalva (1992, 1994). Despite the fact that these units show morphosyntactic properties of both suffixes and words, which makes their classification difficult to establish, we will refer to these elements as suffix(-like) units, following the traditional terminology. In the translation of *z*-evaluative suffixes, we use DIM for diminutive and AUG for augmentative.
275. We use the expression *word stress* to indicate that each element of these constructions bears a word-level stress, instead of a secondary rhythmic stress. This does not preclude, nevertheless, that one of the stressed elements is the prominent element of the whole construction, as we will see further below. For the demonstration that the stress on the base is also assigned with reference to morphological information see, for example, Pereira (1999), who calls this stress “morphological secondary stress”.
276. In Cunha and Cintra (1984) *avos* is written as a separate word. We will adopt this representation, despite the fact that we will assume this element to have the same status as *z*-evaluative suffixes and *–mente*.
277. Inflection is only visible with *z*-evaluative words, since the morphological base in this case may inflect for plural. As for *–mente*, it attaches to adjectival bases that do not inflect for plural, and, although it selects a feminine adjectival base, according to Villalva (1994: 4.2.2.1) gender cannot be considered to be inflectionally marked in EP. Finally, *avos* attaches to numerals that are invariable. Recall, to conclude, that inflection may be marked by a phonetically empty constituent (cf. Villalva 1994, and chapter 5, section 6 of this book).
278. To be exact, the central vowel is not lowered, since it is already low underlyingly. Nevertheless, this vowel patterns like the other vowels in prosodic word final syllables closed by sonorants in that it does not reduce.
279. The example in (348) is adapted from Frota (2000) (the prosodic bracketing is ours). The syllable bearing the focal stress is signaled with capital letters in boldface. Recall that the assignment of focal stress affects I-level prominence and implies the compression of the range immediately after the focal stress (see chapter 3, section 19 for further details). This is an important remark since the first syllable of the word may also bear emphatic stress, which has

- different phonological and semantic/pragmatic consequences (see also chapter 3, section 17).
280. Recall that focal stress is always assigned to the most prominent element of the relevant *word*, which in non-compound words is the syllable bearing word stress.
 281. Andrade (1984: 23) reports that the stressed syllable of the base in words with *z*-evaluative suffixes shows a duration significantly lower than the stressed syllable of the suffix. However, as pointed out by the author, the constructions studied occur in phrase final position. The different values may thus follow from the presence of phrasal prominence on the last prosodic word of the constituent, rather than from the prominence relations within these words.
 282. Notice that this does not preclude that a pitch accent may also appear associated to the first prosodic word of the compound, when the second prosodic word also has a pitch accent.
 283. Two facts point to the relevance of the notion *head of the compound* for the blocking of non-back vowel deletion: (i) the first one is the similarity in the blocking contexts of all processes of vowel deletion – regardless of the specific prosodic domain, vowel deletion always seems to be obligatorily blocked when the target vowel is followed by a vowel belonging to the prominent prosodic word of the constituent it is part of; (ii) the second fact concerns the behavior of vowel deletion in compound words (like those formed by some abbreviations – see section 3.5) that include more than two prosodic words. In the latter cases, vowel deletion is not obligatorily blocked when the second vowel is stressed and belongs to an internal prosodic word, but it is obligatorily blocked when V2 is stressed and belongs to the head of the compound.
 284. This term should not be confused with the term *minimal word*, used to refer to the minimum size of prosodic words in several languages (see chapter 1, section 6.1).
 285. We believe that there is some inter- and intra-speaker variation concerning the possibility of round vowel deletion when two prosodic words belong to the same ϕ but the head of the constituent is not involved. This variation appears to be dependent on register, since vowel deletion is felt to be more likely to occur in more informal registers. What is crucial is that, even in these informal styles, the deletion of the vowel within compounds is totally excluded, and thus a clear distinction emerges between a sequence of two prosodic words within the ϕ -phrase and within the compound prosodic word.
 286. We have not collected data on the possibility of assigning initial stress to these structures, but it seems clear to us that it can also occur in this context, like emphatic stress.
 287. In the prosodic tree, *w* and *s* stand for ‘weak’ and ‘strong’, respectively. As already pointed out, the relevance of prominence relations within the prosodic compound will become clear later on in this chapter.

288. Despite the similarities between prefixes and some stems in morphological compounds, there is at least one test that appears to set these categories apart in EP: non-high non-central vowels in final position of an internal stem of transparent morphological compounds always surface as low (cf. *tél[ɛ]-chamada* ‘lit. distant call’; *mon[ɔ]-acentual* ‘monoaccentual’). Given that this behavior is specific of compound internal stems, examples of reduced final unstressed vowels are expected to be found with prefixes. Although rare, examples of this kind do exist (cf. *vice-*, where the final non-back vowel undergoes vowel reduction and is usually deleted, like similar underlying vowels in prosodic word final position). Interestingly, this morpheme may show a low non-back vowel for some speakers, thus suggesting that these speakers treat it like a stem rather than a prefix.
289. There is dialectal variation with respect to the realization of this type of words. Nevertheless, it is well-known that in the Lisbon variety we are describing [ɐjʃ] is not a natural realization in such non-transparent prefixed words, in contrast with the realization of the prefix in transparent constructions.
290. This apparently differentiates EP from Italian (see chapter 1, section 6.5). According to Nespor and Vogel (1986), monosyllabic prefixes in Italian, in transparent constructions, form prosodic words independent from their bases if they end with a vowel, but they are incorporated into the prosodic word that dominates their base if they end with a consonant. Adopting a different approach, Peperkamp (1997a) claims that in Italian and Spanish monosyllabic prefixes never form independent prosodic words, since they do not observe the minimality requirement assumed to be imposed on prosodic words. In French, by contrast, monosyllabic prefixes are argued to form prosodic words (cf. Hannahs 1995b).
291. It is not possible to test the behavior of these constructions with respect to round vowel deletion and non-back vowel deletion because we could only identify one prefix ending with the relevant vowels (namely, *vice-*) and there are no words where it is followed by a stressed vowel. With respect to syllable degemination, we only found one case where the context for the rule could be met (namely, *vice-secretário* ‘vice-secretary’). We do not think syllable degemination is likely to apply here because the consonants involved are fricatives (but see chapter 3, section 15, for other descriptions according to which syllable degemination is possible in this context). In any event, we judge similarly degemination in this segmental context in non-compound environments (e.g. *disse segredos* ‘(he) told secrets’). Thus, these facts cannot be taken to mean that syllable degemination does not apply within prefixed words formed by two prosodic words. Notice that syllable degemination has been suggested not to occur with prefixes in general (cf. Villalva 1994: 157). Nevertheless, the examples given by Villalva involve either stressless prefixes (which are expected not to undergo the process because

they do not form independent prosodic words), or a segmental string that does not display the context for the rule to apply (as in the example *superperto* ‘very closed’ – see chapter 3, section 15 for the description of this process).

292. As for initial stress and emphatic stress, we have no data concerning the possibilities of assigning these types of stresses to the initial position of the prosodic word that includes the morphological base of stressed prefixes. The same is true for the remaining constructions studied in this chapter. Since our intuitions in these cases are not sound enough, we cannot say whether initial stress and emphatic stress may be assigned to the internal members of compound prosodic words. This is an issue for further research.
293. It is often assumed that the compound linking vowels in the language are –i– and –o– (e.g. Villalva 1994: 6.1.2), the former appearing frequently at the end of the first member of compounds with Latin stems, and the latter occurring often at the end of the first member of compounds with Greek stems (e.g. Cunha and Cintra 1984). From a phonological viewpoint, the linking vowel –o– behaves like non-high non-back vowels and round vowels at the end of internal members of compounds, since all these vowels are realized obligatorily as low (e.g. *israEl*[ɔ]-*palestinIano* ‘israelian-palestinian’, *tÉl*[ɛ]-*chamAda* ‘distant call’, *bI*[ɔ]-*mÉdico* ‘biomedic’, respectively).
294. We should point out that many of the exceptional cases of absence of vowel reduction in stressless positions involve morphological compounds that have lexicalized as a single prosodic word. Examples of exceptionally unreduced stressless vowels are found in words such as *homorgânico* ‘homorganic’ and *homónimo* ‘homonym’, where the first stem is not perceived as stressed and it may count for the computation of the primary word stress. Notice that the initial vowels of these words are obligatorily realized as low, as usually happens with vowels that exceptionally fail to undergo vowel reduction.
295. As the first stems end in [i] or in a low vowel, we may not test the behavior of these compounds with respect to final vowel deletion processes or syllable degemination.
296. With respect to the form *grande área* there seems to be interspeaker variation as to its status as a compound: for some speakers its meaning is compositional, while for other speakers it is merely the name of a part of a football field. For the latter group of speakers, including us, the presence of the glide is obligatory. For the former group of speakers vowel deletion appears to be possible in this context.
297. The difficulty in judging these cases is probably related to the morphosyntactic analysis of this type of expressions. According to Villalva (1994: 6.2.2), speakers hesitate in considering A+N and N+A expressions as compounds given that their interpretation is similar whether they form X° or X^{max}-level expressions, unless they have undergone a process of semantic lexicalization. In the examples under observation, the meaning of the expres-

sions is very close to their compositional meaning. Our hypothesis concerning these structures is that the ambiguity of the morphosyntactic analysis may still have an effect on the prosodization of this type of compounds (which may either form a prosodic compound or not). This issue is left for future investigation.

298. In a reading task with 4 speakers, described in detail in chapter 7, this word was realized with a glide by two speakers (in two realizations – one at normal speech rate and the other at fast speech rate), and with no glide in the two realizations of two other speakers. Although we find the glide possible when it occurs in phrase final position, as it was the case in the text read by the four speakers, we tend to prefer the realization where the vowel is deleted, and more so if the compound word is not in phrase final position, as in (381c). Again, the reasons for this type of variation require further investigation. We should point out, nevertheless, that, while left headed N+N compounds like *cobra-cascavel/cobras-cascavel* ‘rattlesnake(PL)’ allow for inflection in the first noun (cf. Villalva 1994: 381–382), *verde-água* – a left-headed A+N compound – does not seem to allow for inflection at all. Other words denoting colors that have a specific behavior with respect to inflection are *cor de rosa* ‘pink’ and *cor de laranja* ‘orange’. For Villalva (1994: 388), these formations must be considered lexicalized compounds or in a process of lexicalization.
299. According to their behavior with respect to a number of tests considered in Villalva (1994: 6.2.2) (including the (im)possibility of internal inflection, as well as morphological and syntactic modification tests), left headed N+N compounds, and A+A and N+N coordinated compounds are classified as syntactically opaque and morphologically transparent; and expressions composed of N+PP, A+N and N+A are claimed to be both syntactically and morphologically transparent.
300. There is only one letter of the alphabet that starts with a stressless vowel (*h* [ɐgá]), but we found no abbreviations with this letter. In the expression referring to the blood type *Rh+* or *Rh-* ([érɐgá]) there is (nearly) obligatory vowel deletion. This is the expected behavior, given that V2 is unstressed.
301. An issue not treated in this section concerns the maximum number of prosodic words with a compound domain. Frota (2000) suggests that compound intonational phrases are possibly restricted to a binary setting in EP. The data from abbreviations seems to indicate that compound prosodic words may be formed at least by three elements of the same type. In the case of abbreviations formed by four prosodic words, like *PCTP*, *MRPP* and *CNRT* (the former two are abbreviations of Portuguese political parties, and the latter is an abbreviation of an East Timor political party), it is possible to group the first two prosodic words together and then the last two prosodic words. This is shown by the realization of *MRPP* (*eme erre pê pê*, [émjérpépé]), where, for many speakers, including us, the non-back vowel in

the first prosodic word may not be deleted. In this case vowel deletion is obligatorily blocked because the second prosodic word is the head of a prosodic word compound. By contrast, in *CNRT* (*cê ene erre tê*, [séénérté]), the blocking no longer applies. Under our view, this is so because the third prosodic word is not the head of the prosodic word compound that it forms with the following prosodic word. The data points to a binary setting of compound prosodic words in EP, although a ternary setting should also be allowed in structures with an odd number of prosodic words (as, for example, *RFM*).

302. As mentioned in chapter 2, section 4.2, Van der Leeuw (1997) also assumes that pronominal clitics in mesoclitic structures are enclitic to the preceding prosodic word, although no empirical evidence for this analysis is provided.
303. Because the verb in the future and in the conditional forms does not end in a non-back vowel or in a round vowel, we may not assess the behavior of the vowel deletion processes involving these vowels in compound word final position.
304. Notice that an *h* at the beginning of written words in EP is purely conventional, as it never corresponds to an actual speech sound or an underlying segment.
305. The non-independent prosodic word status of the first stem is shown by the fact that its stressed vowel, as well as the linking vowel/the final vowel of the first stem undergoes vowel reduction (to the exception of *autocarro*, where the final vowel of the first stem exceptionally surfaces as low).
306. For example, in EP, it is not clear if DUI obtains with *z*-evaluative suffixed words. In fact, if DUI applies in these cases, the result is ambiguous, i.e. it is not clear whether DUI has applied or only the second coordinate element has a *z*-evaluative suffix. Thus, in the sequence *um café e um chazinho* (a coffee and a tea-DIM) it is not possible to know whether the suffix modifies both members of the coordination or only the second one. This issue is left for future research.
307. We may point out, furthermore, that deletion under identity in EP appears to be restricted to apply to derived words and morphological compounds. In fact, we found no instances of DUI with word+word compounds (e.g. **verde-água e verde azeitona* lit. 'green water and green olive'). With respect to V+N morphosyntactic compounds the following forms could be seen to allow DUI:

- | | | |
|--|---|----------------------------------|
| (i) <i>um abre-latas e abre-garrafas</i> | > | <i>um abre latas e garrafas</i> |
| 'a bottle opener and a can opener' | | |
| <i>um porta-óculos e porta-canetas</i> | > | <i>um porta óculos e canetas</i> |
| 'a glass holder and a pen holder' | | |

However, these cases are probably distinct, since the interpretation here is such that the same object opens cans and bottles / holds glasses and pens. Thus, the elements that are coordinated in these structures appear to be the

second members of the compound, rather than the whole compound words, as in the cases with derived words and morphological compounds. All these issues are addressed in Vigário and Frota (2002), where a full account of DUI in Romance languages is proposed.

308. An interesting pair that clearly shows that the morphosyntactic properties of the original expression are kept in the clipped form is the one presented below: while the form of both clipped words is *super*, in (ia) the word is feminine (as shown by the form of the definite article that precedes it), whereas in (ib) the word is masculine.

- (i) a. *a (revista) Super Interessante* > *a Super*
 the-FEM (magazine) Super Interessante
 b. *o supermercado* > *o super*
 the-MASC supermarket

309. Uttered by a hairdresser.
 310. Uttered by a fireman regularly driving ambulances.
 311. Uttered by Olympic Games athletes during a TV talk show.
 312. Uttered by a nurse.
 313. There are a few cases of clipped words where the EP non-clipped expression does not contain two prosodic words (e.g. *metro* – cf. *metropolitano* ‘subway’; *moto* – cf. *motorizada* ‘motorcycle’). According to Machado (1977), these clipped forms were imported from French. Consequently, they say nothing about the Portuguese rule. Interestingly, unlike in the EP clipped forms based on morphological compounds, where the last vowel of the stem is usually kept low in the clipped word, in these cases the words have been integrated within the system of the language. Thus, in the word *metro*, which is masculine, the final vowel surfaces as high, whereas the word *moto*, which is feminine, has changed into *mota* (the final vowel –*a* corresponds to the usual ending of feminine nouns in the language). Notice, furthermore, that we are excluding from this discussion the shortenings of personal names, whose formation is often specific. For example, while reduplication is not regularly found in other areas of the Portuguese vocabulary, it may occur with proper names (e.g. *Ricardo*>*Cacá*; *Pedro*>*Pepê*; *Jorge*>*Jojó*).
 314. In the work done on other Romance Languages (see the references above), the fact that the resulting clipped words form a disyllabic foot is taken as evidence that the prosodic word in the relevant languages is subject to minimality requirements. As we have seen, the prosodic word in EP does not seem to have a minimal size. It is therefore predicted that clipped forms may be monosyllabic. This prediction is borne out, as shown by the following data: *inter-net* > *net*; *bi-sexual* > *bi*; *e-mail* > *mail*.
 315. Notice, nevertheless, that all function words with more than two syllables are stressed (e.g. *durante* ‘during’, *todavia/embora* ‘although’)
 316. See, for example, Booij (1995: 120) and Peperkamp (1997a) for similar observations in Dutch and Italian, respectively.

317. Following the proposal of Villalva (1994) concerning the theme index and inflection, the derivational stem in nouns and adjectives may be defined as a morphological word minus the theme index and inflection (see also d'Andrade 1988: 110 for a similar definition).
318. To be more accurate, the morphological word node corresponds in Villalva's terms to the maximal morphological projection. For ease of exposition, we use the term *morphological word* with a similar meaning.
319. In chapter 4, section 4.3, Villalva presents a slightly different structure for *-mente* adverbs: *-mente* attaches to the morphological word node, yielding a stem to which the theme index is added, thus forming a theme; finally, inflection is added, thus creating a new morphological word node. For the present proposal, it is irrelevant whether the whole suffix plus theme index and inflection is adjoined to the morphological (word) base, or whether the suffix is added to the morphological word node and forms a new morphological base to which the theme index and inflection are added in turn. In both cases the structure will involve an inner morphological word node (containing, a derivational stem and the inflection constituent), and an outer morphological word node.
320. Recall that the future and conditional may no longer be synchronically related to the structure that gives rise to mesoclis (cf. chapter 4).
321. In the EP literature, the source of word stress in function words and prefixes is in general not discussed (see chapter 2, section 4.1). With respect to prefixes, Pereira (1999) proposes that their stress follows from the fact that they form independent prosodic words. The prosodic word status of these prefixes, in turn, is seen to result from information included in their lexical entry. We assume, by contrast, that prosodic constituency is not stored in the lexical entry of individual items, but rather that it is built according to a number of generalizations and conditions (cf. section 5.3).
322. This well-formedness condition may be seen to follow from principle (d) of the Strict Layer Hypothesis, as formulated in Nespor and Vogel (1986: 7), and presented in (2iv) of this book.
323. Notice that not all compounds display a strong-weak stress pattern in Dutch (cf. Booij 1995; Visch 1999; Nespor 1999b). This pattern alternates with the weak-strong pattern, depending on the lexical category of the members of the compound, as well as on their internal structure. Under the present hypothesis, only when the compound strong-weak pattern is found is the compound word assumed to be mapped onto a compound prosodic word. In the remaining cases, the compound word is not mapped onto a prosodic compound and thus the stress pattern is assumed to depend on the ϕ -level prominence pattern or on lexical rules of stress assignment. According to this hypothesis, a prosodic distinction between compound types is predicted to emerge in Dutch, in a way parallel to EP, where some compounds are prosodized as compound prosodic words, some are grouped directly under ϕ and some are lexicalized.

398 *Notes*

324. A more comprehensive comparison between compounding in EP and other languages is beyond the scope of the present work. We remark, nonetheless, that in Dutch the typical compounds that display the compound stress correspond to the most frequent compound type in the language, that is the nominal compound (cf. Booij 1995: 115).
325. As mentioned in chapter 1, section 6.4, Nespor and Ralli (1996) and Nespor (1999b) explain the prominence relations found in compound words as following from phonological phrase prominence. Nespor (1999b) supports this idea with data of a number of languages where there is a coincidence between the compound stress pattern and the phonological phrase stress pattern (as in Greek, Turkish and several Romance languages). In order to account for the stress pattern of compounds in languages like English or Dutch, further information is admitted to be required, such as the lexical category of the members of the compound.
326. Recall that the proposal that the right boundary of lexical prosodic words is not postlexically projected was driven by the prosodic behavior of pronominal clitics (cf. chapter 5, section 6).
327. Steps (iii) and (iv) seem to overlap. However, not only is (iv) necessary, as already shown in the preceding section, but (iii) is also necessary, as we will see in a few paragraphs.
328. On the arguments that support this proposal, see Villalva (1994: 6.2.2) (see also note 299 of this book for a very brief review).
329. In our terms, unaccented words are lexically unstressed items, which may undergo reduction processes already within the lexical component, and whose reduced forms may thus not be entirely explained by their occurrence in unstressed position postlexically, unlike simple clitics.
330. In Hall's (1999a) data, full vowels are always reduced to schwa in the weak forms, except if they are deleted, which usually happens when they are followed by another vowel. In addition to this rule, German articles are also argued to be subject to a rule that deletes initial consonants (cf. Wiese 1996: 7.4.3).
331. After controlling for a number of variables, Jurafsky, Bell, and Girand (2002) conclude that the differences noticed result from word-predictability factors, which include relative frequency.
332. Notice that, according to Keating this difference disappears when low-frequency content words are considered.
333. On the basis of these results, it is suggested that presence/absence of reduction may be used by speakers to signal new/old information (Bybee 2000: 253).
334. We should point out that the facts just described distinguish vowel reduction in EP from vowel reduction in other languages. For example, according to Van Oostendorp (2000: 131), in English if a vowel can reduce, reduction is always optional. In Dutch, reduction is also usually optional (Booij 1995; Van Oostendorp 2000). Notice that some Dutch vowels always surface

- reduced, and thus are assumed to be underlying schwas. In other cases, unreduced and reduced vowels alternate in morphologically related forms, but reduced forms are obligatory. These cases are assumed by Booij (1995) to result from allomorphy.
335. In an acoustic study on segments deletion in EP, Delgado Martins (1975) reports that [ə] is deleted in 7 out of a total of 442 vowels. Unfortunately, no information is given with respect to the total number of occurrences of this particular vowel or the context of these deletions.
 336. In Brazilian Portuguese, where vowel deletion as a consequence of reduction processes does not occur, there is also vowel deletion due to the resolution of across-words hiatus (e.g. Bisol 1992, 1996; Abaurre 1996).
 337. Again, in Brazilian Portuguese [i] in unstressed positions is also subject to deletion processes conditioned by the segmental context. For example, this vowel like other non-low vowels may be deleted between homorganic consonants (cf. Bisol and Hora 1994).
 338. The phonetic transcription of *devedor* and *despregar* in (440a) are taken from Mateus and d'Andrade (2000: 44); the transcriptions of *professor* and *lambo* are based on acoustic data from Vigário (1998a: Appendix II) and on perception data from Mateus and Delgado Martins (1982), respectively.
 339. According to E. d'Andrade, a personal communication cited in Andrade (1980: 67), [i] truncation is less favored after a word initial consonant (which is the only word initial context where schwa may occur). This may be seen to follow from the prominence properties of the prosodic word initial position. Recall that this position may be assigned initial and emphatic stress, it defines the context for the application of initial /r/-strengthening, and vowels in prosodic word initial position do not undergo full (lexical) vowel reduction (see in particular chapter 3, section 10).
 340. Notice, however, that this analysis was abandoned in more recent work (e.g. Mateus and d'Andrade 2000).
 341. Repetitions were allowed and all natural-sounding renditions were considered for analysis.
 342. Complementizers in intonational phrase final position, which surface with a schwa, were eliminated from this analysis. A few renditions read with hesitations were also not considered.
 343. The importance of vowel quality and stress in the following vowel for schwa deletion is suggested in Gonçalves Viana (1883: 39–40). However, our observations did not confirm the relevance of these variables.
 344. For this preliminary observation we have considered the data of two of the three speakers whose data is analyzed in the present chapter. The data of another speaker was also analyzed. Because this subject had many hesitations in the reading task and she became unavailable for the recording of the second set of materials we have decided to eliminate her renditions. We

- should point out, nevertheless, that the eliminated data showed precisely the same tendencies as the data described below (see Vigário 1998b).
345. In a study conducted simultaneously, Frota (2000: 5.2.3.1) reports the same tendency for the function word *aos* (to-the-MASC-PL) ([awʃ]/[ɔʃ]) to reduce less in intonational phrase initial position when compared to I-internal position. On the reduction of this function word, see also our results in section 4.2 below.
 346. No noticeable differences were found in the renditions of the different texts. This does not mean that register is not relevant for schwa deletion or other reduction phenomena. Rather, we believe that, in general, speakers have probably employed a formal register during the reading sessions. This is supported by the inspection of the productions of other function words in the task of retelling the tale. Some reduced forms that intuitively are style dependent, like [pə] (<*para*) and [ku] (<*com o*), occurred predominantly in this task.
 347. We have eliminated from the data analyzed a few cases where the position of clitic function words could not be undoubtedly ascertained due to unclear cases of intonational phrase boundaries.
 348. The reason for this may be stylistic. For some speakers the unreduced form of *pelo(s)/pela(s)* is felt to belong to rather marked registers. The existence of the nominal form *pelo* [pélu] 'hair' may also contribute for the preference for the reduced form of the function word.
 349. We should point out that the deletion of the vowel of the preposition, which may also occur with the article *a*, is usually felt to belong to very marked registers. As these examples show, however, it also occurs in other types of speech, though possibly less frequently. We should add that in 3 tokens realized by the same speaker a denasalized form of *com* was found followed by a full word. The number of cases of this sort, however, seems negligible and does not undermine the generalization that reduced forms of *com* are typically found when followed by articles.
 350. In this respect, speakers seem to behave differently. IS has never realized the relevant vowel as a glide when the following vowel is stressed. We also tend to follow this pattern. We should point out that, although the presence of the glide in this context is reminiscent to the presence of the glide in stressed environments within compound prosodic words (cf. chapter 6), the two cases are not identical: while the glide is obligatory when the following vowel bears the compound prosodic word prominence, in this case, the absence of the glide is possible, occurring in 57% of the realizations of RI and IV where the relevant context obtains.
 351. Recall that the list of Bacelar, Marques, and Segura da Cruz (1987) includes the most frequent words of EP, based on a total of ca. 700 000 collected words.

352. The word initial fricative may also be deleted in the inflected forms of the verb *estar* ‘to be’, possibly with the exception of the forms of the future and conditional, where deletion seems a bit more marked.
353. In the more reduced form we believe only the meaning ‘great’ is possible. We should add that this form is also subject to a process that consists of the change of a schwa-vowel (or sometimes [u]) into [ə], which is usually associated with very marked/young people’s registers and typically occurs in imperative or exclamative sentences, as in the examples below:

<i>Que grande parvo!</i>	[kigrɛ̃dɐpárvu]/[kɐgɛ̃dɐpárvu]	‘What a great jerk!’
<i>Vou-te contar!...</i>	[vótɕõtár]/[vótɛkõtár]	‘I’m telling you!...’
<i>Ele é muito giro.</i>	[élémũjtɕĩru]/[élémũjtɛzĩru]	‘He is very cute.’
<i>Jesus!</i>	[ʒizú]/[ʒɛzú]	‘Jesus Christ!’

354. Word final fricative deletion also occurs with the clitic *mas* [mɐʃ] ‘but’ when followed by *não* ‘not’.
355. Our contention is not that reduction may not affect these words, but rather that these forms are judged to be ill-formed, contrary to those in (463) and (464a–468a). Whether the contrast noticed follows from allomorphy or from the fact that reduction may simply be more frequent in the words in (463) is a question we leave open (see also the discussion in section 5.5). Notice, additionally, that the observed differences could be thought to be related to the content versus functional nature of the relevant words. However, most items in (469) are function elements. Furthermore, Jurafsky et al. (2001) show for English that content words may also be subject to different amounts of reduction, depending on their relative frequency.
356. The negative word *não* [ˈnẽw] ‘not’ may also show different realization, in this case depending on the segmental context. For example, the negative sentences presented in Vigário (1998a) (see in particular Appendix II), include sequences such as *não emprestaram* ‘(they) did not lend’ and *não ofereceram* ‘(they) did not offer’, which are in general transcribed as [nẽwɛpʃtárɐw] and [nẽwɔfsérɐw]; however, the transcription [nẽɕpʃtárɐw] and [nõɔfsérɐw] is also found (e.g. sentences 33 and 32 of Appendix II, respectively) (in these transcriptions, primary word stress was only marked in words with more than one syllable). Notice that in all such cases, the negative word was realized with a stress that makes the whole word to become perceptively highlighted. Thus, here we have evidence that the emergence of these forms is not a consequence of destressing. The prosodic status of negative words, and in particular of the sentence negation particle, requires further investigation, which will be left for future work.
357. Diachronic data provide additional examples of possible reductions occurring in similar environments. For example, in the evolution of the form of Latin pronouns that originated the definite articles and accusative pronouns

o(s)/a(s) the-MASC(PL)/FEM(PL) of Portuguese, the pronoun's initial syllable has been deleted in the whole paradigm. This is shown in (i), taken from Williams (1975).

(i) Classic Latin	Vulgar Latin (reconstructed)
<i>illum</i> (accusative, singular)	<i>lo</i> (masculine, singular in Portuguese)
<i>illam</i> (accusative, singular)	<i>la</i> (feminine, singular in Portuguese)
<i>illos</i> (accusative, plural)	<i>los</i> (masculine, plural in Portuguese)
<i>illas</i> (accusative, plural)	<i>las</i> (feminine, plural in Portuguese)

The items *no(s)/na(s)* in-the-MASC(PL)/FEM(PL), formerly, *eno(s)/ena(s)* (cf. Williams 1975: §137.4) illustrate the same point.

358. In her discussion on the clitic status of the Italian pronoun *loro*, Nespor (1999a: 875) explicitly raises and leaves open the question of whether disyllabicity may be used as a diagnostic against the phonological clitic status of a word. That is, clitics are conjectured to be at most monosyllabic universally, as suggested by Neijt (1985) and Berendsen (1986: chap. 2, note 2), among others. In EP, there is evidence that at least four disyllabic words are clitic – *para* ‘for’, *pelo(s)/pela(s)* by-the-MASC(PL)/FEM(PL), *porque* ‘because’ and *cada* ‘each’. Nonetheless, EP follows the general tendency in that the great majority of clitic words are at most monosyllabic. We should add, furthermore, that the reduction of disyllabic clitics via the deletion of the vowel of the first syllable results in the creation of a complex onset. Interestingly, in the case of *para* the complex onset eventually becomes simple through consonant deletion.
359. In accordance with our analysis of proclitic function words (see chapter 5), we are assuming that stress in the proclitic counts as a prosodic word initial stress. Since such stress may only occur once per prosodic word, whether or not with prosodic adjuncts, the number of potential prosodic word initial positions relevant for initial stress is identical to the number of prosodic words in the corpus. Prosodic words starting with a syllable bearing word stress and not preceded by a proclitic are, naturally, excluded.
360. According to Martins (1994: chaps.3–4), postposition of clitics is the dominant pattern in the 13th century, whereas preverbal clitic location becomes almost generalized in the 15th century. Nevertheless, free variation in clitic positioning is still attested in 16th century.
361. The same general observation is made in Galves and Galves (1995) and Britto (1998), although the shift is considered to be more recent. The precise point in time where this change took place is not relevant for our discussion.
362. On the phonology of enclisis and proclisis in Old French see also Jacobs (1993).
363. The relevance of prosodic information in the distribution of clitics is also proposed in Halpern (1995), Taylor (1996), and Pintzuk (1996) for languages such as Serbo-Croatian, Bulgarian, Old French, Ancient Greek, Old English,

among others. According to Booij (1996a), the Dutch pronoun *ie*, unlike other clitics in the language, also requires a prosodic host to its left, and thus may not occur sentence initially (see chapter 1, section 6.3 of this book).

364. Notice furthermore that three sentences of the BP corpus were excluded (see the next note), and thus the PB corpus is slightly shorter than the EP one. According to Frota and Vigário (2000), the perceived prominences in PB follow an alternating strong-week pattern. Additionally, unlike in EP, in BP prosodic word internal non-primary stresses are very frequent, occurring in 69.7% of the cases where the context for the alternating pattern obtains.
365. Three sentences read with hesitations were excluded from the materials.
366. For the purposes of this argumentation we are assuming, with Galves and Galves (1995), that the relevant prosodic properties of BP are similar to those of Classical Portuguese spoken before the change took place in the European variety. Other prosodic and rhythmic differences between the two varieties of Portuguese, some of which bearing on the phonological properties of the intonational phrase, are described in Frota and Vigário (2000, 2001).
367. There are a few cases where both the lexicalized fused form and the regular non-lexicalized sequence may occur, as illustrated below (see also section 5.5):

<i>em um, em uma</i>	<i>num, numa</i>	in a-MASC, in a-FEM
<i>em outro, em outra</i>	<i>noutro, noutra</i>	in other-MASC, in other-FEM
<i>em algum, em alguma</i>	<i>nalgum, nalguma</i>	in some-MASC, in some-FEM

368. *pelo(s)/pela(s)* have lexicalized in a period when the form of the preposition that became *por* was *per* and the form of the article after a word ending in consonant was *lo(s)/la(s)*. The deletion of the preposition final consonant is a process similar to the one that gave rise to the precompiled rule involving verbs ending with consonant when followed by the accusative pronoun with the form *lo(s)/la(s)* (see chapter 4, section 4.2.2). On the history of these forms, see Williams (1975: §143).
369. Notice, nonetheless, that sequences of pronominal clitics may also lexicalize, as we have seen in chapter 4, section 4.2.2. We should recall that pronominal clitics behave like prepositions in showing a rate of reduction higher than complementizers.
370. Fusion with *com* has in fact occurred in personal pronouns, as shown in (i) below. This case is nevertheless slightly different, since the forms *migo, tigo, nosco, vosco*, which were formed in Latin (*me, te, nos, vos* plus *cum*) had lost their original meaning and were not used as independent words in other contexts when fusion took place, unlike in all the cases reported above.

(i) <i>comigo</i>	‘with me’	<i>connosco</i>	‘with us’
<i>contigo</i>	‘with you(SG)’	<i>convosco</i>	‘with you(PL)’

371. In our corpus, we only have 4 cases of *sem*. In all the realizations of these cases the three speakers, this word was pronounced as [sẽj]. We believe, in

fact, that the reduction of this function word is totally impossible. We should add that, although *em* is much more frequent than *sem*, according to the EP frequency list in Bacelar, Marques, and Segura da Cruz (1987), *sem* is among the 140 most frequent words of EP. It is conceivable that the reason why this word does not undergo reduction is because it bears word stress, unlike *em*, but we do not have independent evidence in favour of this hypothesis at the present moment.

372. In fact, the same conditions for lexicalization may also obtain with *que* in expressions such as *se bem que* ‘although’ (lit. ‘if well that’), *ainda que* ‘even though’ (lit. ‘still that’), *já que* ‘since’ (lit. ‘already that’), *só que* ‘although’ (lit. ‘only that’). In a few cases, when such expressions occur in intonational phrase final position, *que* may surface without stress, similar to what obligatorily happens with pronominal enclitics (see chapter 5, section 4.3.1). Since this behavior is exceptional, it may be seen to follow from lexical information. Notice, nevertheless, that the exceptional form involving the enclisis of *que* must coexist with the regular form, since this complementizer may also behave like other proclitic words (see also section 5.5 for the evidence that lexically stored allomorphs may be in free variation with non-lexicalized units).
373. This difference is noticed by EP speakers, which often judge the reduced forms of consonant-schwa complementizers to denote a very marked register, unlike in the case of reduced prepositions. The different behavior of the two categories may already be inferred from the description of Gonçalves Viana (1883).
374. In order to have comparable data, we have excluded from the calculation of Bacelar, Marques, and Segura da Cruz data the cases where *de* has lexicalized with the following function word.
375. We should point out, additionally, that *de* may also function as a complementizer and thus some of the cases listed in Bacelar, Marques, and Segura da Cruz (1987) under this form are probably complementizers. The actual number of occurrences of *de* functioning as a preposition is therefore likely to be slightly lower. Notice further that *que* may also function as an interrogative word (both in the frequency corpus of EP, and in the corpus we have analyzed). Its occurrence as a complementizer is, nevertheless, much more common in the language.
376. In Frota and Vigário (1996) and Vigário and Frota (1998) a hypothesis is put forward that the set of items that function as proclisis triggers in EP is the set of strong function words. The facts reported in this chapter, and in previous work (Vigário 1999a), add to this idea, since they show that complementizers behave like elements stronger than prepositions, in that they reduce less and often appear in I-prominent positions. Recall that complementizers, but not prepositions, are proclisis triggers in EP.

377. It should be noticed, nevertheless, that precompiled allomorphy can also be hypothesized to be dependent on phrasal prosodic information (e.g. Nespor 1990) and on speech register.
378. Notice that the realization of the glide is a consequence of the application of a general process that affects the first unstressed vowel of a VV sequence, which is V1-semivocalization (cf. chapter 3, section 11.2).
379. The same observation applies to other English function words, as Jurafsky, Bell, and Girand (2002) list 82 common pronunciations of the 10 most frequent function words of English (including 9 variants of the article *the*).
380. For example, the reduced form [prə] of the clitic *para* 'to' is attested in our corpus in I-final stressed position. I-initial stress in this form was also occasionally signaled.
381. Unfortunately, we did not systematically investigate the assignment of prosodic word or intonational phrase initial prominence in our corpus, nor did we transcribe the stressed forms that occurred in I-final position. Therefore, we cannot determine with precision which clitic forms may receive such stresses. The identification of the clitic forms that may bear stress, which we propose should not follow from phonetic reduction, is thus left for future research.
382. We believe that this proposal may also be extended to BP. The importance of morphological information for the prosodization of prefixes may explain why, according to Schwindt (2000), synchronic (non-lexicalized) prefixes in BP are prosodized similar to EP (cf. Vigário 1999a; see also chapter 5, section 3.2). As for the prosodization of proclitics in BP, the data in Schwindt (2000) and Bisol (2000) clearly indicate that, like prefixes, proclitics are not incorporated into the prosodic word that dominates their base/host. Nevertheless, the identification of the type of phonological host and the prosodic relation between proclitics and host remains unclear. These issues are not systematically addressed in Schwindt's work because he adopts the view that clitics are part of the clitic group. Bisol (2000) is less categorical in this regard, admitting that the clitic group may be reinterpreted as a postlexical prosodic word.
383. We should observe, nonetheless, that word-level stress has consequences in the definition of the (lexical) prosodic word in EP similar to phonotactic constraints in Dutch and German. That is why we have proposed in chapter 6, section 5.2 that not only morphosyntactic but also purely phonological information is required for the construction of this prosodic domain in EP.
384. We assume that resyllabification induces prosodic word restructuring, along the lines of Peperkamp (1997a) (see chapter 1, section 7).
385. To the best of our knowledge, EP does not present any phenomenon that undoubtedly refers to the foot. With respect to word stress, its location appears to be obtained to a great extent from lexical and morphological information (see Mateus 1983 and the literature reviewed in chapter 2 and chapter 3, section 2). Except for word-stress – if at all –, it is not clear what role this

- constituent may play in the phonology of the language (see also chapter 5, section 4.3.2, for a discussion on secondary stress in EP).
386. In EP this type of neutralization affects other vowels, as well, and applies to unstressed vowels within the word, regardless of their location with respect to word stress.
 387. Although compounds are in general not referred to display a phonological behaviour distinct from phonological phrases in BP, Lee (1995: 3.2) suggests that stress-clash resolution does distinguish between lexical compounds and sequences of words combined postlexically, in that it may apply to the former but not to the latter. He suggests, furthermore, that stress-clash resolution in BP is a lexical phenomenon. However, clash resolution in this variety is reported in other studies to apply optionally within the phonological phrase (cf., among others, Major 1985: 4.2; Abousalh 1997). According to L. Bisol (personal communication), at least in some varieties of BP stress clash resolution is obligatory within compounds but optional within phonological phrases.
 388. Kleinhenz (1997) proposes that some languages are word-based and some are phrase-based, in the sense that the prosodic phonology of languages is organized on the basis of the prosodic word or the phonological phrase, respectively. EP is given as an example of a word-based language, whereas French and Chi Mwi:ni represent phrase-based languages. German and Dutch are proposed to be intermediate stage languages. The argument for the classification of EP as word-based is founded on Frota's (1996) observation that ϕ is in general invisible in EP. The data gathered in this book can now be added in order to support the view that EP phonology extensively refers to the prosodic word. It should be noticed, nevertheless, that phonological phenomena that refer to ϕ , although rare, do exist in EP (see chapter 1, §1.1.2, and note 389 below).
 389. We have seen that there is language variation in the role that prosodic constituents play in particular phonological systems. What the consequences of this type of crosslinguistic variation for language acquisition may be is too soon to tell. We may observe, nevertheless, that work on EP has shown that ϕ is a weak prosodic domain (e.g. no sandhi phenomena refer to this constituent, it is often not tonally marked, and frequently ϕ -level prominence is not perceived – cf. Frota 2000; Vigário 1998a). Despite this, children do not seem to have any difficulty in acquiring the word order of the language. Interestingly, the (scarce) evidence that has been gathered for ϕ in EP consist of phenomena that somehow refer to the stress or the head of this constituent (see chapter 1, section 2). Whether this may indicate that there is a limit to phonological constituents' weakness in those domains that play a major role in the acquisition of grammatical information, is a question we leave open.

Personal copy

References

- Abaurre, M. Bernardete
 1996 Acento frasal e processos fonológicos segmentais. In: Leda Bisol (ed.), *Letras de Hoje* 31(2): 41–50.
- Abousalh, Elaine S. Ferreira
 1997 Resolução de choques acentuais de acento no português brasileiro: elementos para uma reflexão sobre a interface sintaxe-fonologia. MA Dissertation, University of Campinas.
- Adams, Marianne
 1989 Verb second effects in Medieval French. In: Carl Kirschner and Janet Decesaris (eds.), *Studies in Romance Linguistics*, 1–31. Amsterdam: John Benjamins.
- Âmbar, Manuela
 1992 *Para uma sintaxe da inversão sujeito-verbo em Português*. Lisboa: Colibri.
- Anderson, Stephen
 1992 *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- Andrade, Amália
 1980 Vowel Reduction and Related Phenomena in Portuguese. MA Dissertation, University of Leeds.
- Andrade, Amália
 1984 Acoustic study of vowel duration in European Portuguese (based on one subject). *Relatórios do Grupo de Fonética e Fonologia* nº 5. Lisboa: INIC.
- Araújo, Gabriel
 2001 Processos morfológicos marginais no Português Brasileiro: Truncamento e reduplicação. Paper given at the *XVII Encontro Nacional da Associação Portuguesa de Linguística*, Lisbon.
- Bacelar, M. Fernanda, M. Lúcia Marques and M. Luísa Segura da Cruz
 1987 *Português Fundamental. Vol. II. Métodos e Documentos, Tomo 1: Inquérito de Frequência*. Lisboa: INIC/CLUL.
- Bafile, Laura
 1994 La riassegnazione postlessicale dell'accento nel napoletano. *Quaderni del Dipartimento di Linguistica - Università di Firenze* 5: 1–23.
- Bafile, Laura
 1997 Parole grammaticali e struttura prosodica: dati dell'italiano e del napoletano. *Lingua e Stile* a. XXXII, 3: 433–469.

408 *References*

- Barbosa, Pilar
1996 Clitic Placement in European Portuguese and the Position of Subjects. In: Aaron L. Halpern and Arnold M. Zwicky (eds.), *Approaching Second. Second Position Clitics and Related Phenomena*, 1–40. Stanford: CSLI Publications.
- Bauer, Laurie
1987 *English Word-Formation*. Cambridge: Cambridge University Press (3rd ed.).
- Baumann, Monika
1996 (Re)syllabification Across Word Boundaries: Psycholinguistic Evidence from Dutch Clitics. In: Artemis Alexiadou, Nanna Fuhrhop, Paul Law and Sylvia Löhken (eds.), *ZAS Papers in Linguistics* 6: 28–42.
- Berendsen, Egon
1985 Tracing Case in Phonology. *Natural Language and Linguistic Theory* 3: 95–106.
- Berendsen, Egon
1986 *The Phonology of Cliticization*. Dordrecht: Foris.
- Bisol, Leda
1992 Sândi vocálico externo: degeminação e elisão. In: M. Bernardete Abaurre and Leo Wetzels (eds.), *Cadernos de Estudos Lingüísticos* 23: 83–101.
- Bisol, Leda
1996 O sândi e a ressilabificação. In: Leda Bisol (ed.), *Letras de Hoje* 31(2): 159–168.
- Bisol, Leda
2000 O Clítico e o seu Status Prosódico. *Revista de Estudos de Linguagem UFMG* 9(1): 5–30.
- Bisol, Leda and Dermeval da Hora
1994 Palatalização da oclusiva dental e fonologia lexical. In: *Actas do IX Encontro Nacional da Associação Portuguesa de Linguística*, 61–80. Lisboa: APL/Colobri.
- Booij, Geert
1985 Coordination Reduction in Complex Words: A case for prosodic phonology. In: Harry van der Hulst and Norval Smith (eds.), *Advances in Nonlinear Phonology*, 143–160. Dordrecht: Foris.
- Booij, Geert
1988 On the relation between lexical and prosodic phonology. In: Pier Marco Bertinetto and Michele Loporcaro (eds.), *Certamen Phonologicum*, 63–76. Torino: Rosenberg and Sellier.
- Booij, Geert
1994 Lexical Phonology: a Review. *Lingua e Stile* a. XXIX, 4: 525–555.
- Booij, Geert
1995 *The Phonology of Dutch*. Oxford: Clarendon Press.

- Booij, Geert
1996a Cliticization as prosodic integration: the case of Dutch. *The Linguistic Review* 13: 219–242.
- Booij, Geert
1996b Lexical Phonology and the derivational residue. In: Jacques Durand and Bernard Laks (eds.), *Current Trends in Phonology: Models and Methods*, 69–96. Salford: European Studies Research Institute.
- Booij, Geert
1999 The Role of the Prosodic Word in Phonotactic Generalizations. In: T. Alan Hall and Ursula Kleinhenz (eds.), *Studies on the Phonological Word*, 47–72. Amsterdam: John Benjamins.
- Booij, Geert and Jerzy Rubach
1984 Morphological and prosodic domains in Lexical Phonology. *Phonology Yearbook* 1: 1–29.
- Booij, Geert and Jerzy Rubach
1987 Postcyclic versus Postlexical Rules in Lexical Phonology. *Linguistic Inquiry* 18: 1–44.
- Booij, Geert and Rochelle Lieber
1993 On the simultaneity of morphological and prosodic structure. In: Sharon Hargus and Ellen Kaisse (eds.), *Studies in Lexical Phonology*, 23–44. San Diego: Academic Press.
- Brandão de Carvalho, Joaquim
1988 Nasalité et structure syllabique en Portugais et en galicien: approche linéaire et panchronique d'un problème phonologique. *Verba* 15: 237–263.
- Brandão de Carvalho, Joaquim
1989 Phonological conditions on Portuguese clitic placement: on syntactic evidence for stress and rhythmical patterns. *Linguistics* 27: 405–436.
- Brandão de Carvalho, Joaquim
1994 What are vowels made of? The 'no-rule' approach and particle phonology. *Studia Linguistica* 48: 1–27.
- Britto, Helena
1998 Clitic Placement in the History of European Portuguese: 17th and 18th Centuries. Paper given at the *Workshop on Rhythmic Patterns, Parameter Setting and Language Change*. December, IEL, Universidade de Campinas.
- Bybee, Joan
1985 *Morphology*. Amsterdam: John Benjamins.
- Bybee, Joan
2000 Lexicalization of sound change and alternating environments. In: Michael B. Broe and Janet B. Pierrehumbert (eds.), *Papers in Laboratory Phonology. Acquisition and the Lexicon*, 250–268. Cambridge: Cambridge University Press.

410 *References*

- Cabré, Teresa
 1993 *Estrutura Gramatical i Lexico: el Mot Minim en Catalana*. Ph.D. Dissertation, Universitat Autònoma de Barcelona.
- Cabré, Teresa and Michael Kenstowicz
 1995 Prosodic Trapping in Catalan. *Linguistic Inquiry* 26: 694–705.
- Câmara, Joaquim Mattoso Jr.
 1953 *Para o estudo da fonêmica Portuguesa*. Organização Simões: Rio de Janeiro.
- Câmara, Joaquim Mattoso Jr.
 1972 *The Portuguese Language*. Chicago: University of Chicago Press.
- Cambier-Langeveld, Tina
 1997. The domain of final lengthening in the production of Dutch. In: J. Coerts and H. de Hoop (eds.), *Linguistics in the Netherlands 1997*, 13–24. Amsterdam, Philadelphia: John Benjamins.
- Cambier-Langeveld, Tina
 2000 *Temporal Marking of Accents and Boundaries*. The Hague: Holland Academic Graphics.
- Cameira, Célia
 1994 Da intuição à teoria – Questões sintático-fonológicas do clítico – o. In: *Actas do IX Encontro da Associação Portuguesa de Linguística*, 99–112. Lisboa: APL/Colibri.
- Carstairs-McCarty, Andrew
 1998. Phonological constraints on Morphological Rules. In: Andrew Spencer and Arnold Zwicky (eds.), *The Handbook of Morphology*, 144–148. Oxford: Blackwell Publishers.
- Chomsky, Noam
 1986 *Barriers*. Cambridge, MA.: MIT Press.
- Chomsky, Noam
 1995 *The Minimalist Program*. Cambridge, Mass.: The MIT Press.
- Chomsky, Noam and Morris Halle
 1968 *The Sound Pattern of English*. New York: Harper and Row.
- Clements, G.N.
 1978 Tone and syntax in Ewe. In: Donna Jo Napoli (ed.), *Elements of Tone, Stress, and Intonation*, 21–99. Washington, D.C.: Georgetown University Press.
- Clements, G.N.
 1988 Towards a substantive theory of features specification. In: J. Blevins and J. Carter (eds.), *Proceedings of NELS 18*, 79–93. Amherst: GSLA.
- Clements, G.N. and Elizabeth Hume
 1995 Internal organization of speech sounds. In: John Goldsmith (ed.), *The Handbook of Phonological Theory*, 245–306. Cambridge, Mass.: Basil Blackwell.

- Collischonn, Gisela
1999 O acento em Português. In: Leda Bisol (ed.), *Introdução a Estudos de Fonologia do Português Brasileiro*, 125–158. Porto Alegre: EDIPUCRS (2nd ed.–revised).
- Condoravdi, Cleo
1990 Sandhi Rules of Greek and Prosodic Theory. In: Sharon Inkelas and Draga Zec (eds.), *The Phonology-Syntax Connection*, 63–84. Chicago: University of Chicago Press.
- Cowper, Elizabeth A. and Keren D. Rice
1987 Are phonosyntactic rules necessary? *Phonology Yearbook* 4: 185–194.
- Crysmann, Berthold
1999 Licensing Proclisis in European Portuguese. In: F. Corblin, C. Dobrovie-Sorin and J.M. Marandin (eds.), *Empirical Issues in Formal Syntax and Semantics 2. Selected Proceedings from the Colloque de Syntaxe et Sémantique de Paris*, 255–276. Thesus: The Hague.
- Crysmann, Berthold
2000a Clitics and Coordination in Linear Structure. In: B. Gerlach and J. Grijzenhout (eds.), *Clitics in Phonology, Morphology, and Syntax*, 121–159. Linguistics Today. John Benjamins.
- Crysmann, Berthold
2000b Syntactic Transparency of Pronominal Affixes. In: Ronnie Cann, Claire Grover and Philip Miller (eds.), *Grammatical Interfaces in HPSG*, 77–96. Stanford: CSLI Publications.
- Cunha, Celso and Luís Filipe Lindley Cintra
1984 *Nova Gramática do Português Contemporâneo*. Lisboa: Edições Sá da Costa (2nd ed.).
- d'Andrade, Ernesto
1977 *Aspects de la phonologie (générative) du Portugais*. Lisboa: INIC.
- d'Andrade, Ernesto
1988 O acento de Palavra em Português. In: J. Staczek (ed.), *On Spanish, Portuguese and Catalan Linguistics*, 17–37. Washington: Georgetown University Press.
- d'Andrade, Ernesto
1994a Uma mudança fonética. In: *Temas de Fonologia*, 33–40. Lisboa: Colibri.
- d'Andrade, Ernesto
1994b Uma Palavra, um Plural. In: *Temas de Fonologia*, 41–46. Lisboa: Colibri.
- d'Andrade, Ernesto
1994c Histórias de O. In: *Temas de Fonologia*, 47–57. Lisboa: Colibri.

412 *References*

- d' Andrade, Ernesto
1994d As alternâncias vocálicas no sistema verbal do Português. In: *Temas de Fonologia*, 95–106. Lisboa: Colibri.
- d' Andrade, Ernesto
1993 *Dicionário Inverso do Português*. Lisboa: Edições Cosmos.
- d' Andrade, Ernesto
1997 Some remarks about stress in Portuguese. In: Fernando Martínez-Gil and Alfonso Morales-Front (eds.), *Phonology and Morphology of the Major Iberian Languages*, 343–358. Washington, D.C.: Georgetown University Press.
- d' Andrade, Ernesto
1998. Sobre a alternância vogal/glide em Português. In: M. Antónia Mota and Rita Marquilhas (eds.), *Actas do XIII Encontro Nacional da Associação Portuguesa de Linguística*, 91–102. Lisboa: APL/Colibri.
- d' Andrade, Ernesto and Alain Kihm
1988 Fonologia auto-segmental e nasais em Português. In: *Actas do III Encontro da Associação Portuguesa de Linguística*, 51–60. Lisboa: APL.
- d' Andrade, Ernesto and Bernard Laks
1992 Na crista da onda: o acento de palavra em português. In: *Actas do VII Encontro Nacional da Associação Portuguesa de Linguística*, 15–26. Lisboa: APL/Colibri.
- d' Andrade, Ernesto and M. Celeste Rodrigues
1999 Das escolas e das culturas: história de uma sequência consonântica. In: Ana Cristina Macário Lopes and Cristina Martins (eds.), *Actas do XIV Encontro da Associação Portuguesa de Linguística*, 117–133. Braga: APL.
- d' Andrade, Ernesto and M. Céu Viana
1989 Ainda sobre o acento e o ritmo em Português. In: *Actas do IV Encontro da Associação Portuguesa de Linguística*, 3–16. Lisboa: APL.
- d' Andrade, Ernesto and M. Céu Viana
1993 Que horas são às (1)3 e 15? In: *Actas do VIII Encontro da Associação Portuguesa de Linguística*, 59–66. Lisboa: APL.
- d' Andrade, Ernesto and M. Céu Viana
1994 Sinérese, diérese e estrutura silábica. In: *Actas do IX Encontro da Associação Portuguesa de Linguística*, 31–42. Lisboa: APL/Colibri.
- d' Andrade, Ernesto and M. Céu Viana
1999 Constantino e os acidentes de Constantinopla: os acentos do Português e do Castelhana. In: Palmira Marrafa and M. Antónia Mota (eds.), *Linguística Computacional. Investigação Fundamental e Aplicações*, 79–95. Lisboa: APL/Colibri.

- Delgado Martins, M. Raquel
1975 Vogais e consoantes do Português: estatística de ocorrência, duração e intensidade. *Boletim de Filologia*, 24(1): 1–11.
- Dell, François
1985 L'accentuation des phrases en Français. In: François Dell, Daniel Hirst and Jean-Roger Vergnaud (eds.), *Forme sonore du langage*, 65–122. Paris: Hermann.
- Di Sciullo, Anna-Maria and Edwin Williams
1987 *On the Definition of Word*. Cambridge, Mass.: MIT Press.
- Lello, José and Edgar Lello
1964 *Dicionário Prático Ilustrado. Novo Dicionário Enciclopédico Luso-Brasileiro*. Porto: Lello & Irmãos Editores.
- Dicionário Universal da Língua Portuguesa*
1995 Lisboa: Texto Editora.
- Dixon, Robert M.W.
1977 Some Phonological Rules of Yedjn. *Linguistic Inquiry* 8: 1–34.
- Downing, Laura J.
1999. Prosodic stem≠prosodic word in Bantu. In: T. Alan Hall and Ursula Kleinhenz (eds.), *Studies on the Phonological Word*, 73–98. Amsterdam: John Benjamins.
- Dresher, B. Elan
1996 Introduction to Metrical and Prosodic Phonology. In: James L. Morgan and Katherine Demuth (eds.), *Signal to Syntax: Bootstrapping From Speech to Grammar in Early Acquisition*, 41–54. Mahwah, NJ: Lawrence Erlbaum Associates.
- Duarte, Inês and Gabriela Matos
2000 Romance Clitics and the Minimalist Program. In: João Costa (ed.), *Portuguese Syntax. New Comparative Studies*, 116–142. Oxford: Oxford University Press.
- Duarte, Inês, Gabriela Matos and Isabel Faria
1995 Specificity of European Portuguese clitics in Romance. In: Isabel Faria and M. João Freitas (eds.), *Studies on the Acquisition of Portuguese*, 129–154. Lisboa: APL/Colibri.
- Ellison, Mark and M. Céu Viana
1996 Antagonismo e elisão das vogais átonas em P.E. In: Inês Duarte e Matilde Miguel (eds.), *Actas do XI Encontro Nacional da Associação Portuguesa de Linguística*, 261–282. Volume 3. Lisboa: APL/Colibri.
- Falé, Isabel
1995 Fragmento da Prosódia do Português Europeu: as Estruturas Coordenadas. MA Dissertation, University of Lisbon.

414 *References*

- Fitzpatrick-Cole, Jennifer
 1991 The Minimal Word in Bengali. In: Aaron Halpern (ed.), *The Proceedings of the Ninth West Coast Conference on Formal Linguistics*, 157–170. Stanford: Stanford Linguistics Association.
- Fowler, C. and J. Housum
 1987 Talkers' signaling of 'new' and 'old' words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language* 26: 489–504.
- Freitas, M. João
 1995 Alveolar trill(ions of problems): Evidence from children acquiring European Portuguese syllables. In: Isabel Faria and M. João Freitas (eds.), *Studies on the Acquisition of Portuguese*, 55–68. Lisboa: APL/Colibri.
- Freitas, M. João
 1997 Aquisição da Estrutura Silábica do Português Europeu. Ph.D. Dissertation, University of Lisbon.
- Frota, Sónia
 1991 Para a Prosódia da Frase: Quantificador, Advérbio e Marcação Prosódica (somente alguns tópicos em foco). MA Dissertation, University of Lisbon.
- Frota, Sónia
 1994a Is Focus a phonological category in Portuguese? In: Peter Ackema and Maaïke Schoorlemmer (eds.), *Proceedings of ConSole 1*, 69–86. The Hague: Holland Academic Graphics.
- Frota, Sónia
 1994b On the Prosody of Focus in European Portuguese. *Proceedings of the Workshop on Phonology*, 45–66. Lisboa: APL.
- Frota, Sónia
 1995 Clashes and Prosodic Domains in EP. *Proceedings 19*, 93–107. Institute of Phonetic Sciences, University of Amsterdam.
- Frota, Sónia
 1996 Prosodic phrases and European Portuguese: in search of evidence. In: Antonietta Bisetti, Laura Brugè, João Costa, Rob Goedemans, Nicola Munaro and Ruben van de Vijver (eds.), *Proceedings of ConSole 3*, 47–69. Leiden: SOLE.
- Frota, Sónia
 1997 On the prosody and intonation of Focus in European Portuguese. In: Fernando Martínez-Gil and Alfonso Morales-Front (eds.), *Phonology and Morphology of the Major Iberian Languages*, 359–392. Washington, D.C.: Georgetown University Press.
- Frota, Sónia
 2000 *Prosody and Focus in European Portuguese. Phonological Phrasing and Intonation*. New York: Garland Publishing.

- Frota, Sónia
2002a Nuclear falls and rises in European Portuguese: a phonological analysis of declarative and question intonation. *Probus* 14: 113–146.
- Frota, Sónia
2002b Tonal association and target alignment in European Portuguese nuclear falls. In: Carlos Gussenhoven and Natasha Warner (eds.), *Laboratory Phonology* 7, 387–418. Berlin: Mouton de Gruyter.
- Frota, Sónia and Marina Vigário
1996 On Weight Effects in European Portuguese. Paper given at the *Glow Workshop On Weight Effects*, Athens.
- Frota, Sónia and Marina Vigário
2000 Aspectos de prosódia comparada: ritmo e entoação no PE e no PB. In: Rui Vieira de Castro and Pilar Barbosa (eds.), *Actas do XV Encontro da Associação Portuguesa de Linguística*, 533–555. Volume 1. Coimbra: APL.
- Frota, Sónia and Marina Vigário
2001 On the correlates of rhythmic distinctions: The European/Brazilian Portuguese case. *Probus* 13: 247–275.
- Frota, Sónia and Marina Vigário
2002 Efeitos de peso no Português Europeu. In: Maria Helena Mateus and Clara Nunes Correia (eds.), *Saberes no Tempo. Homenagem a Henriqueta Costa Campos*, 315–333. Lisboa: Edições Colibri.
- Fulmer, S. Lee
1991 A Case of Inflection before Derivation. In: Dawn Bates (ed.), *Proceedings of the Tenth West Coast Conference on Formal Linguistics*, 151–162. Stanford: Stanford Linguistics Association.
- Galves, António and Charlotte Galves
1995 A Case Study of Prosody Driven Language Change. University of S. Paulo / University of Campinas, ms.
- Galves, Charlotte
1992 Clitic placement in European Portuguese, evidence for a non-homogeneous theory of enclisis. In: *Actas do Workshop sobre o Português*, 61–80. Lisboa: APL.
- Galves, Charlotte
2000 Prosody, clitics and language change in Portuguese. Paper given at the workshop *Statistical Physics, Pattern Identification and Language Change*, Lisbon.
- Godjevac, Svetlana M.A.
2000 Intonation, word order, and focus projection in Serbo-Croatian. Ph.D. Dissertation, Ohio State University.
- Goldsmith, John
1990 *Autosegmental and Metrical Phonology*. Cambridge, Mass.: Basil Blackwell.

416 *References*

- Gonçalves Viana, A.R.
 1883 Essais de phonétique et de phonologie de la langue portugaise d'après le dialect actuel de Lisbonne. In: *Estudos de Fonética Portuguesa*, 83–152. 1973. Lisboa: Imprensa Nacional - Casa da Moeda.
- Gonçalves, Anabela
 1994 Aspectos da reestruturação sintáctica em Português Europeu: as construções com verbos modais. In: *Actas do IX Encontro da Associação Portuguesa de Linguística*, 235–250. Lisboa: APL/Colibri.
- Gonçalves, Anabela
 1999 *Predicados complexos verbais em contextos de infinitivo não-preposicionado do Português Europeu*. Ph.D. Dissertation, University of Lisbon.
- Green, Anthony Dubach.
 1997 *The Prosodic Structure of Irish, Scots Gaelic, and Manx*. Ph.D. Dissertation, Cornell University.
- Greenberg, S., D. Ellis, and J. Hollenback
 1996 Insights into spoken language gleaned from phonetic transcription of the Switchboard corpus. In: *Proceedings of the 1996 International Conference on Spoken Language Processing*, 24–27.
- Gussenhoven, Carlos and Toni Rietveld
 1992 Intonation contours, prosodic structure and preboundary lengthening. *Journal of Phonetics* 20: 283–303.
- Hall, T. Alan
 1994 Final Extraprosodicity and Resyllabification. In: Richard Wiese (ed.), *Recent Developments in Lexical Phonology*, 151–166. Düsseldorf: Heinrich Heine Universität.
- Hall, T. Alan
 1999a. Phonotactics and the Prosodic Structure of German Function Words. In: T. Alan Hall and Ursula Kleihenz (eds.), *Studies on the Phonological Word*, 99–131. Amsterdam: John Benjamins.
- Hall, T. Alan
 1999b The Phonological Word: a Review. In: T. Alan Hall and Ursula Kleinhenz (eds.), *Studies on the Phonological Word*, 1–22. Amsterdam: John Benjamins.
- Halpern, Aaron
 1991 Phonological Evidence for the Surface Nonexistence of Empty Categories. In: *Proceedings of the Ninth West Coast Conference on Formal Linguistics*, 233–247. Stanford: Stanford Linguistics Association.
- Halpern, Aaron
 1995 *On the Placement and Morphology of Clitics*. Stanford: CSLI Publications.

- Halpern, Aaron
1996 Cluster morphology. In: Wolfgang U. Dressler (ed.), *Advances in Morphology*, 51–61. New York: Mouton de Gruyter.
- Halpern, Aaron
1998 Clitics. In: Andrew Spencer and Arnold M. Zwicky (eds.), *The Handbook of Morphology*, 101–122. Oxford: Blackwell Publishers.
- Hannahs, Stephen J.
1995a *Prosodic Structure and French Morphophonology*. Tübingen: Max Niemeyer Verlag.
- Hannahs, Stephen J.
1995b Glide formation, prefixation, and the phonological word in French. In: Jon Amastae, Grant Goodall, Mario Montalbetti and Marianne Phinney (eds.), *Contemporary Research in Romance Linguistics*, 13–24. Amsterdam/Philadelphia: John Benjamins Publishers.
- Harris, James W.
1983 *Syllable Structure and Stress in Spanish*. Cambridge, MA: MIT Press.
- Hauge, Kjetil Rå
1976 The Word Order of Predicate Clitics in Bulgarian. Universitetet i Oslo. Slavisk-Baltisk Institutt, Oslo (Meddelelser, 10).
- Hayes, Bruce
1989 The prosodic hierarchy in meter. In: Paul Kiparsky and Gilbert Youmans (eds.), *Rhythm and Meter. Phonetics and Phonology 1*, 201–260. New York: Academic Press.
- Hayes, Bruce
1990 Precompiled Phrasal Phonology. In: Sharon Inkelas and Draga Zec (eds.), *The Phonology-Syntax Connection*, 85–108. Chicago: University of Chicago Press.
- Hayes, Bruce and Aditi Lahiri
1991 Bengali Intonational Phonology. *Natural Language and Linguistic Theory* 9: 47–96.
- Hulst, Harry van der
1996 Separating primary and secondary accent. In: Rob Goedemans, Harry van der Hulst and Ellis Visch (eds.), *Stress Patterns of the World. Part I: Background*, 1–25. The Hague: Holland Academic Graphics.
- Hulst, Harry van der, Bernadet Hendriks and Jeroen van de Weijer
1999 A survey of word prosodic systems of European languages. In: Harry van der Hulst (ed.), *Word Prosodic Systems in the Languages of Europe*, 425–575. Berlin: Mouton de Gruyter.
- Idsardi, William
1992 The Computation of Prosody. Ph.D. Dissertation, MIT.

418 *References*

- Inkelas, Sharon
1990 *Prosodic Constituency in the Lexicon*. New York: Garland Publishing.
- Itô, Junko
1986 *Syllable Theory in Prosodic Phonology*. Ph.D. Dissertation, University of Massachusetts.
- Itô, Junko and Armin Mester
1992 *Weak Layering and Word Binarity*. Ms., University of California, Santa Cruz.
- Jackendoff, Ray
1972 *Semantic Interpretation in Generative Grammar*. Cambridge, Mass.: MIT Press.
- Jacobs, Haike
1993 The Phonology of Enclisis and Proclisis in Gallo-Romance and Old French. In: Ashby et al. (eds.), *Linguistic Perspectives on the Romance Languages*, 149–163. Amsterdam / Philadelphia: John Benjamins.
- Jun, Sun-Ah
1996 *The Phonetics and Phonology of Korean Prosody: Intonational Phonology and Prosodic Structure*. New York: Garland Publishing.
- Jurafsky, Daniel, Alan Bell, Michelle Gregory and William D. Raymond
2001 Probabilistic Relations between Words: Evidence from Reduction in Lexical Production. In: Joan Bybee and Paul Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*, 229–254. Amsterdam: John Benjamins.
- Jurafsky, Daniel, Alan Bell and Cynthia Girard
2002 The role of lemma in form variation. In: Carlos Gussenhoven and Natasha Warner (eds.), *Laboratory Phonology 7*, 3–34. Berlin: Mouton de Gruyter.
- Kaiser, Georg A.
1999 A evolução dos pronomes clíticos no Português Europeu em comparação com os pronomes clíticos em outras línguas Românicas. In: José Antonio S. Padilla and Magnolia Troya Déniz (eds.), *Actas del XI Congreso Internacional de la Asociación de Lingüística y Filología de la América Latina, 1933–1943*. Volume III. Gran Canaria: Universidad de Las Palmas.
- Kaisse, Ellen
1985 *Connected Speech. The Interaction of Syntax and Phonology*. San Diego: Academic Press.
- Kaisse, Ellen
1990 Towards a Typology of Postlexical Rules. In: Sharon Inkelas and Draga Zec (eds.), *The Phonology-Syntax Connection*, 127–143. Chicago: The University of Chicago Press.

- Kaisse, Ellen and Patricia Shaw
 1985 On the theory of lexical phonology. *Phonology Yearbook* 2: 1–30.
- Keating, Patricia A.
 1998 Word-level phonetic variation in large speech corpora. In: Artemis Alexiadou, Nanna Fuhrhop, Ursula Kleinhenz and Paul Law (eds.), *ZAS Papers in Linguistics 11. Papers of the Conference on 'The word as a phonetic unit'*, 35–50.
- Kenstowicz, Michael
 1994 *Phonology in Generative Grammar*. Cambridge, MA: Blackwell.
- Kiparsky, Paul
 1982 From cyclic phonology to lexical phonology. In: Harry van der Hulst and Norval Smith (eds.), *The Structure of Phonological Representations*, 131–175. Volume I. Dordrecht: Foris.
- Kiparsky, Paul
 1985 Some consequences of Lexical Phonology. *Phonology Yearbook* 2: 85–138.
- Klavans, Judith L.
 1985 The independence of syntax and phonology in cliticization. *Language* 61: 95–120.
- Kleinhenz, Ursula
 1994 *Focus and Phrasing in German*. Arbeitspapiere de Sonderforschungsbereichs 340.
- Kleinhenz, Ursula
 1996 The Prosody of German Clitics. In: Artemis Alexiadou, Nanna Fuhrhop, Paul Law and Sylvia Löhken (eds.), *ZAS Papers in Linguistics* 6, 81–95.
- Kleinhenz, Ursula
 1997 Domain Typology at the Phonology-Syntax Interface. In: Gabriela Matos, Matilde Miguel, Inês Duarte and Isabel Faria (eds.), *Interfaces in Linguistic Theory*, 201–220. Lisboa: APL/Colibri.
- Kraft, C.H. and A.H.M. Kirk-Greene
 1973 *Hausa*. London: Holder and Stoughton.
- Ladd, D. Robert
 1992 Compound Prosodic Domains. University of Edinburgh, ms.
- Ladd, D. Robert
 1996 *Intonational Phonology*. Cambridge: Cambridge University Press.
- Leben, William and Firmin Ahoua
 1997 Prosodic domains in Baule. *Phonology* 14: 113–132.
- Lee, Seung-Hwa
 1995 Morfologia e Fonologia Lexical do Português do Brasil. Ph.D. Dissertation, University of Campinas.
- Leeuw, Frank van der
 1997 *Clitics. Prosodic Studies*. (HIL Dissertations 29.) The Hague: Holland Academic Graphics.

420 *References*

- Lëfstedt, Ingvar
1992 Swedish segment length and structure preservation. *Studia Linguistica* 46(2): 93–127.
- Leite de Vasconcellos, José
1901 *Esquisse d'une Dialectologie Portugaise*. 3rd ed. 1987. Lisboa: INIC/CLUL.
- Lexicoteca. Novo Dicionário da Língua Portuguesa*
1984 Lisboa: Círculo de Leitores.
- Liberman, Mark and Alan Prince
1977 On stress and linguistic rhythm. *Linguistic Inquiry* 8: 249–336.
- Lloret, Maria-Rosa
2002 Estrutura silàbica. In: Joan Solà, Maria-Rosa Lloret, Joan Mascaró and Manuel Pérez Saldanya (eds.), *Gramàtica del Català Contemporani*, 195–249. Volume I. Barcelona: Editorial Empúries.
- Lüdtke, Helmut
1953 Fonemática portuguesa. *Boletim de Filologia* 14: 197–217.
- Machado, José Pedro
1977 *Dicionário Etimológico da Língua Portuguesa*. Lisboa: Livros Horizonte (3rd ed).
- Madeira, Ana Maria
1992 On clitic placement in European Portuguese. *UCL Working Papers in Linguistics* 4, 95–122.
- Major, Roy C.
1985 Stress and rhythm in Brazilian Portuguese. *Language* 61(2): 259–282.
- Manzini, Rita
1998 Syntactic approaches to cliticization. *Glott International*, 3(3), 3–7.
- Marquilha, Rita
2000 A elevação do vocalismo átono pretónico em Português Europeu: nova cronologia. Paper given at the *Congresso Internacional “500 anos da Língua Portuguesa no Brasil”*, Évora.
- Martins, Ana Maria
1994 Clíticos na História do Português. Ph.D. Dissertation, University of Lisbon.
- Marquilha, Rita
2000 A elevação do vocalismo átono pretónico em Português Europeu: nova cronologia. Paper given at *Congresso Internacional “500 anos da Língua Portuguesa no Brasil”*, Évora.
- Mascarenhas, Isabel
1997 Estudo da Variação Dialectal entre Lisboa e Porto das Vogais átonas [-rec] e [+arr] em Contexto Inicial. MA Dissertation, University of Lisbon.

- Mateus, Maria Helena
1975 *Aspectos da Fonologia Portuguesa*. 2nd ed.–revised, 1983. Lisboa: INIC.
- Mateus, Maria Helena
1983 O acento de palavra em Português: uma nova proposta. *Boletim de Filologia* 28: 211–229.
- Mateus, Maria Helena
1995 A silabificação de base em Português. In: *Actas X Encontro Nacional da Associação Portuguesa de Linguística*, 289–300. Lisboa: Colibri/APL.
- Mateus, Maria Helena
1997a Aspectos da fonologia lexical do português. In: Ana Maria Brito, Fátima Oliveira, Isabel Pires de Lima and Rosa M. Martelo (eds.), *Sentido que a vida faz. Estudos para Óscar Lopes*, 693–703. Porto: Campo das Letras.
- Mateus, Maria Helena
1997b Redundâncias lexicais e subespecificação: o sistema do português. In: Ivo Castro (ed.), *Actas do XII Encontro da Associação Portuguesa de Linguística*, 203–213. Volume 1. Lisboa: APL/Colibri.
- Mateus, Maria Helena and Ernesto d' Andrade
1998 The syllable structure in European Portuguese. *D.E.L.T.A.* 14(1): 13–32.
- Mateus, Maria Helena and Ernesto d' Andrade
2000 *The Phonology of Portuguese*. Oxford: Oxford University Press.
- Mateus, Maria Helena and M. Raquel Delgado Martins
1982 Contribuição para o estudo das vogais átonas [ə] e [u] no Português Europeu. *Biblos* LVIII: 111–128.
- Mateus, Maria Helena, Amália Andrade, M. Céu Viana and Alina Villalva
1990 *Fonética, Fonologia e Morfologia do Português*. Lisboa: Universidade Aberta.
- Matos, Gabriela
1997 Configurações sintáticas em estruturas de colocação simultânea de clítico. In: Ana Maria Brito, Fátima Oliveira, Isabel Pires de Lima and Rosa M. Martelo (eds.), *Sentido que a vida faz. Estudos para Óscar Lopes*, 705–717. Porto: Campo das Letras.
- Mazuca, Reiko
1996 Can a Grammatical Parameter Be Set Before the First Word? In: James L. Morgan and Katherine Demuth (eds.), *Signal to Syntax: Bootstrapping From Speech to Grammar in Early Acquisition*, 313–330. Mahwah, NJ: Lawrence Erlbaum Associates.
- Mester, R.A.
1990 Patterns of Truncation. *Linguistic Inquiry* 21: 478–485.

422 *References*

- Miller, H. Philip and Ivan A. Sag
 1997 French clitic movement without clitics or movement. *Natural Language and Linguistic Theory* 15: 573–639.
- Mohanan, K. P.
 1986 *The Theory of Lexical Phonology*. Dordrecht: D. Reidel Publ. Company.
- Monachesi, Paula
 1996 On the Representation of Italian Clitics. In: Ursula Kleinhenz (ed.), *Interfaces in Phonology*, 83–101. (Studia Grammatica 41.) Berlin: Akademie Verlag.
- Moraes, João Antônio
 1997. Intonation in Brazilian Portuguese. In: Daniel Hirst and Albert Di Cristo (eds.), *Intonation Systems. A Survey of Twenty Languages*, 179–194. Cambridge: Cambridge University Press.
- Morais Barbosa, Jorge
 1965 *Etudes de phonologie portugaise*. Lisboa: Junta de Investigações do Ultramar [2nd ed., 1983. Évora: Universidade de Évora].
- Morais Barbosa, Jorge
 1988 Notas sobre a pronúncia portuguesa nos últimos cem anos. *Biblos* LXIV: 329–382.
- Morales-Front, Alfonso and Eric Holt
 1997 On the interplay of morphology, prosody and faithfulness in Portuguese pluralization. In: Fernando Martínez-Gil and Alfonso Morales-Front (eds.), *Phonology and Morphology of the Major Iberian Languages*, 393–437. Washington, D.C.: Georgetown University Press.
- Moreno, Cláudio
 1997 Morfologia nominal do português: um estudo de fonologia lexical. Ph.D. Dissertation, Pontifícia Universidade Católica do Rio Grande do Sul.
- Morgan, James L., Rushen Shi and Paul Allopenna
 1996 Perceptual Basis of Rudimentary Grammatical Categories: Toward a Broader Conceptualization of Bootstrapping. In: James L Morgan and Katherine Demuth (eds.), *Signal to Syntax: Bootstrapping From Speech to Grammar in Early Acquisition*, 263–283. Mahwah, NJ: Lawrence Erlbaum Associates.
- Mota, M. Antónia
 2000 Mudanças anunciadas em Português: esboço de comparação entre o Português Europeu e o Português Brasileiro. Paper given at the International Conference *500 anos da Língua Portuguesa no Brasil*. Évora, May.
- Napoli, Donna Jo and Joel Nevis
 1987 Inflected prepositions in Italian. *Phonology Yearbook* 4: 195–209.

- Neijt, Anneke
1985 Clitics in arboreal phonology. In: Harry van der Hulst and Norval Smith (eds.), *Advances in Nonlinear Phonology*, 179–192. Dordrecht: Foris.
- Nespor, Marina
1985 The Phonological Word in Italian. In: Harry van der Hulst and Norval Smith (eds.), *Advances in Nonlinear Phonology*, 193–204. Dordrecht: Foris.
- Nespor, Marina
1986 The phonological word in Greek and Italian. In: Henning Andersen (ed.), *Sandhi Phenomena in the Languages of Europe*, 65–74. Berlin: Mouton de Gruyter.
- Nespor, Marina
1990 Vowel deletion in Italian: the organization of the phonological component. *The Linguistic Review* 7: 375–398.
- Nespor, Marina
1999a The phonology of clitic groups. In: Henk van Riemsdijk (ed.), *Clitics in the Languages of Europe*, 865–887. Berlin: Mouton de Gruyter.
- Nespor, Marina
1999b Stress Domains. In: Harry van der Hulst (ed.), *Word Prosodic Systems in the Languages of Europe*, 117–159. Berlin: Mouton de Gruyter.
- Nespor, Marina and Angela Ralli
1996 Morphology-phonology interface: Phonological domains in Greek compounds. *The Linguistic Review* 13: 357–382.
- Nespor, Marina and Irene Vogel
1982 Prosodic domains of external sandhi rules. In: Harry van der Hulst and Norval Smith (eds.), *The Structure of Phonological Representations*, 225–255. Volume I. Dordrecht: Foris.
- Nespor, Marina and Irene Vogel
1983 Prosodic structure above the word. In: Anne Cutler and D. Robert Ladd (eds.), *Prosody: Models and Measurements*, 123–140. Berlin: Springer.
- Nespor, Marina and Irene Vogel
1986 *Prosodic Phonology*. Dordrecht: Foris
- Nespor, Marina and Irene Vogel
1989 On clashes and lapses. *Phonology* 6: 69–116.
- Nespor, Marina and M. Scorretti
1985 Empty Elements and Phonological Form. In: J. Guéron, H.-G. Obenauer and J.Y. Pollock (eds.), *Grammatical Representations*, 203–235. Dordrecht: Foris.

424 *References*

- Nespor, Marina, Maria Teresa Guasti and Anne Christophe
 1996 Selecting Word Order: The Rhythmic Activation Principle. In: Ursula Kleinhenz (ed.), *Interfaces in Phonology*, 1–26. Berlin: Akademie Verlag.
- Novo Michaelis. Dicionário Ilustrado*
 1983 São Paulo: Melhoramentos (30th ed.).
- Noyer, Rolf
 1994 Mobile Affixes in Huave: Optimality and Morphological Wellformedness. In: Erin Duncan, Donka Farkas and Philip Spaelti (eds.), *Proceedings of the Twelfth West Coast Conference on Formal Linguistics*, 67–82. Stanford: Stanford Linguistics Association.
- Odden, David
 1990 Syntax, Lexical Rules and Postlexical Rules in Kimatumbi. In: Sharon Inkelas and Draga Zec (eds.), *The Phonology-Syntax Connection*, 259–277. Chicago: The University of Chicago Press.
- Oostendorp, Mark van
 2000 *Phonological Projection - A Theory of Feature Content and Prosodic Structure*. Berlin: Mouton de Gruyter.
- Parker, Steve
 1999 On the behavior of definite articles in Chamicuro. *Language* 75(3): 552–562.
- Parkinson, Stephen
 1988 Portuguese. In: Martin Harris and Nigel Vincent (eds.), *The Romance Languages*, 131–169. London: Routledge.
- Peperkamp, Sharon
 1996 On the prosodic representation of clitics. In: Ursula Kleinhenz (ed.), *Interfaces in Phonology*, 102–127. (Studia Grammatica 41.) Berlin: Akademie Verlag.
- Peperkamp, Sharon
 1997a *Prosodic Words*. HIL Dissertations 34. The Hague: Holland Academic Graphics.
- Peperkamp, Sharon
 1997b The prosodic structure of compounds. In: Gabriela Matos, Matilde Miguel, Inês Duarte and Isabel Faria (eds.), *Interfaces in Linguistic Theory*, 259–279. Lisboa: APL/Colibri.
- Pereira, Isabel
 1990 Da Prosódia: Análise da Evolução do Conceito de Prosódia e das Diferentes Abordagens Linguísticas das Questões Prosódicas. MA Dissertation, University of Lisbon.
- Pereira, Isabel
 1999 O Acento de Palavra em Português. Uma Análise Métrica. Ph.D. Dissertation, University of Lisbon.
- Pesetsky, David
 1985 Morphology and logical form. *Linguistic Inquiry* 16: 193–248.

- Peters, Ann M. and Sven Strömquist
 1996 The Role of Prosody in the Acquisition of Grammatical Morphemes. In: James L. Morgan and Katherine Demuth (eds.), *Signal to Syntax: Bootstrapping From Speech to Grammar in Early Acquisition*, 215–232. Mahwah, NJ: Lawrence Erlbaum Associates.
- Pierrehumbert, Janet and Mary Beckman
 1988 *Japanese Tone Structure*. Cambridge, Mass.: MIT Press.
- Pintzuk, Susan
 1996 Cliticization in Old English. In: Aaron Halpern and Arnold M. Zwicky (eds.), *Approaching Second. Second Position Clitics and Related Phenomena*, 375–409. Stanford: CSLI Publications.
- Prieto, Pilar
 1992 Truncation processes in Spanish. *Studies in the Linguistic Sciences* 22: 143–158.
- Prince, Alan and Paul Smolensky
 1993 Optimality Theory: Constraint Interaction in Generative Grammar. Rutgers University and University of Colorado, ms.
- Pulgram, E.
 1970 *Syllable, Word, Nexus, Cursus, Series Minor 81*. The Hague: Mouton de Gruyter.
- Raffelsiefen, Renate
 1999a Diagnostics for Prosodic Words Revisited: The Case of Historically Prefixed Words in English. In: T. Alan Hall and Ursula Kleinhenz (eds.), *Studies on the Phonological Word*, 133–201. Amsterdam: John Benjamins.
- Raffelsiefen, Renate
 1999b Phonological constraints on English word formation. In: Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1998*, 225–287. Dordrecht: Kluwer Academic Publishers.
- Ramus, Frank, Marina Nespor and Jacques Mehler
 1999 Correlates of linguistic rhythm in speech. *Cognition* 73: 265–292.
- Reighard, J. and A. de Almeida
 1983 /R/ en début et en fin syllabiques dans les dialectes portugaises. *Revue de l'Association Québécoise de Linguistique* 3: 177–186.
- Rice, D. Keren
 1989 On eliminating Resyllabification into Onsets. In: E. Jane Fee and Katherine Hunt (eds.), *Proceedings of the Eighth West Coast Conference on Formal Linguistics*, 331–346. Stanford: Stanford Linguistics Association.
- Roca, Iggy M.
 1986 Secondary stress and metrical rhythm. *Phonology Yearbook* 3: 341–370.

426 *References*

- Roca, Iggy M.
1999 Stress in the Romance Languages. In: Harry van der Hulst (ed.), *Word Prosodic Systems in the Languages of Europe*, 659–811. Berlin: Mouton de Gruyter.
- Rochemont, Michael and Peter Culicover
1990 *English focus constructions and the theory of grammar*. Cambridge: Cambridge University Press.
- Rodrigues, M. Celeste
2001 *Lisboa e Braga: variação e fonologia*. Ph.D. Dissertation, Lisbon University.
- Rouveret, Alain
1999 Clitics, subjects and tense in European Portuguese. In: Henk van Riemsdijk (eds.), *Clitics in the Languages of Europe*, 639–677. Berlin: Mouton de Gruyter.
- Rudin, Catherine, Christina Kramer, Loren Billings and Matthew Baerman
1999 Macedonian and Bulgarian *li* Questions: Beyond Syntax. *Natural Language and Linguistic Theory* 17: 541–586.
- Sá Nogueira, Rodrigo
1938 *Elementos para um tratado de fonética portuguesa*. Lisboa: Centro de Estudos Filológicos.
- Said Ali, Manuel
1964 *Gramática Secundária da Língua Portuguesa*. São Paulo: Edições Melhoramentos (5th ed.).
- Scalise, Sergio
1984 *Generative Morphology*. Dordrecht: Foris.
- Schwindt, Luíz Carlos
2000 O prefixo no Português Brasileiro: análise morfofonológica. Ph.D. Dissertation, Pontifícia Universidade Católica do Rio Grande do Sul.
- Selkirk, Elisabeth
1972 The Phrase Phonology of English and French. Ph.D. dissertation, MIT. [Published in 1981 by the Indiana University Linguistics Club, Bloomington, Indiana.]
- Selkirk, Elisabeth
1980 Prosodic Domains in Phonology. In: M. Aronoff and M.-L. Kean (eds.), *Juncture*, 107–129. Saratoga, Calif.: Anma Libri.
- Selkirk, Elisabeth
1984 *Phonology and Syntax. The Relation between Sound and Structure*. Cambridge, MA: MIT Press.
- Selkirk, Elisabeth
1986 On derived domains in sentence phonology. *Phonology Yearbook* 3: 371–405.
- Selkirk, Elisabeth
1996 The prosodic structure of function words. In: James L. Morgan and Katherine Demuth (eds.), *Signal to Syntax: Bootstrapping From*

- Speech to Grammar in Early Acquisition*, 187–213. Mahwah, NJ: Lawrence Erlbaum Associates. [An earlier version was published in: Jill N. Beckman, Laura W. Dickey and Susane Urbanczyk (eds.) (1995), *Papers in Optimality Theory*, 439–469. University of Massachusetts Occasional Papers 18. Amherest, MA: Graduate Linguistic Student Association.]
- Skorge, Sílvia
1956 Os sufixos diminutivos em Português. *Boletim de Filologia* 16(1–2): 50–90. Lisboa: Centro de Estudos Filológicos.
- Spencer, Andrew
1991 *Morphological Theory*. Oxford: Blackwell.
- Sportiche, Dominique
1992 Clitic constructions. UCLA, ms..
- Sproat, Richard
1988 Bracketing Paradoxes, Cliticization and Other Topics: The Mapping Between Syntactic and Phonological Structures. In: Martin Everaert, Arnold Evers, Riny Huybregts and Mieke Trommelen (eds.), *Morphology and Modularity*, 339–360. Dordrecht: Foris Publications.
- Statistica*
1997 General Conventions & Statistics I. USA: StatSoft (<http://www.statsoft.com>).
- Taylor, Ann
1996 A Prosodic Account of Clitic Position in Ancient Greek. In: Aaron Halpern and Arnold M. Zwicky (eds.), *Approaching Second. Second Position Clitics and Related Phenomena*, 477–503. Stanford: CSLI Publications.
- Teyssier, Paul
1980 *Histoire de la Langue Portugaise*. Presses Universitaires de France. [Portuguese translation: *História da Língua Portuguesa*. 2nd ed. 1984. Lisboa: Livraria Sá da Costa.]
- Thornton, Anna M.
1996 On Some Phenomena of Prosodic Morphology in Italian: Accorciamenti, Hipocoristics and Prosodic Delemitation. *Probus* 8: 81–112.
- Truckenbrodt, Hubert
1999 On the Relation between Syntactic Phrases and Phonological Phrases. *Linguistic Inquiry* 30: 219–255.
- Venditti, Jennifer J., Sun-Ah Jun, Mary E. Beckman
1996 Prosodic Cues to Syntactic and Other Linguistic Structures in Japanese, Korean, and English. In: James L. Morgan and Katherine Demuth (eds.), *Signal to Syntax: Bootstrapping From Speech to Grammar in Early Acquisition*, 287–311. Mahwah, NJ: Lawrence Erlbaum Associates.

428 *References*

- Verluyten, P.
1982 Recherches sur la prosodie et la métrique du français. Ph.D. Dissertation, University of Antwerp.
- Viana, M. Céu
1987 Para a Síntese da Entoação do Português. Dissertação em Linguística Portuguesa para acesso à categoria de Investigador Auxiliar, CLUL/INIC.
- Vigário, Marina
1997a Marcação prosódica em frases negativas no Português Europeu. In: Ivo Castro (ed.), *Actas do XII Encontro da Associação Portuguesa de Linguística*, 329–349. Volume I. Lisboa: APL/Colibri.
- Vigário, Marina
1997b Processos de desambiguação prosódica em estruturas com advérbios de exclusão. In: Ana Maria Brito, Fátima Oliveira, Isabel Pires de Lima and Rosa M. Martelo (eds.), *Sentido que a vida faz. Estudos para Óscar Lopes*, 855–868. Porto: Campo das Letras.
- Vigário, Marina
1998a *Aspectos da Prosódia do Português Europeu: estruturas com Advérbios de Exclusão e Negação Frásica*. Braga: Universidade do Minho/CEHUM.
- Vigário, Marina
1998b Elisão de Vogal Não-recuada Final e a Palavra Prosódica no Português Europeu. In: *Actas do XIII Encontro Nacional da Associação Portuguesa de Linguística*, 359–376. Lisboa: APL/Colibri.
- Vigário, Marina
1998c The prosodization of clitic function words in European Portuguese. In: Tina Cambier-Langeveld, Anikó Lipták and Michael Redford (eds.), *Proceedings of ConSole VI*, 253–267. Leiden: Sole.
- Vigário, Marina
1999a On the Prosodic Status of Stressless Function Words in European Portuguese. In: T. Alan Hall and Ursula Kleinhenz (eds.), *Studies on the Phonological Word*, 255–294. Amsterdam: John Benjamins.
- Vigário, Marina
1999b Cliticização pronominal no Português Europeu: uma operação pós-lexical. In: *Actas do XIV Encontro da Associação Portuguesa de Linguística*, 577–598. Braga: APL.
- Vigário, Marina
1999c Redução das palavras funcionais clíticas no Português Europeu. Paper given at the conference *Diferenças Rítmicas entre o Português de Portugal e o Português do Brasil*, Universidade Estadual de Campinas.

- Vigário, Marina
 1999e Pronominal cliticization in European Portuguese: a postlexical operation. *Catalan Working Papers in Linguistics* 7: 219–237.
- Vigário, Marina
 2000a The phonology of clitics in Modern European Portuguese. Paper given at the workshop *Statistical Physics, Pattern Identification and Language Change*, Lisbon.
- Vigário, Marina
 2000b Palavra prosódica e composição no Português Europeu. In: Rui Vieira de Castro and Pilar Barbosa (eds.), *Actas do XV Encontro da Associação Portuguesa de Linguística*, 583–602. Volume 2. Coimbra: APL.
- Vigário, Marina and Isabel Falé
 1994 A Sílabas no Português Fundamental: uma descrição e algumas considerações de ordem teórica. In: *Actas do IX Encontro da Associação Portuguesa de Linguística*, 465–477. Lisboa: APL/Colibri.
- Vigário, Marina and Sónia Frota
 1998 Between syntax and phonology: on phrasal weight effects in European Portuguese. Paper given at the 8th *Colloquium on Generative Grammar*, Palmela.
- Vigário, Marina and Sónia Frota
 2000 Contribution to the working session “The syntax-phonology interface”, at the *Extended Workshop Statistical Physics, Pattern Identification and Language Change*, Lisbon.
- Vigário, Marina and Sónia Frota
 2002 Prosodic word deletion in coordinated structures and the syntax-phonology interface. To appear in *Journal of Portuguese Linguistics* 1(2).
- Villalva, Alina
 1992 Portuguese Compounds. *Rivista di Linguistica* 4.1: 201–219.
- Villava, Alina
 1994 Estruturas Morfológicas. Unidades e Hierarquias nas Palavras do Português. Ph.D. Dissertation, University of Lisbon
- Visch, Ellis
 1999 The rhythmic organization of compounds and phrases. In: Harry van der Hulst (ed.), *Word Prosodic Systems in the Languages of Europe*, 161–231. Berlin: Mouton de Gruyter.
- Vogel, Irene
 1990 The clitic group in prosodic phonology. In: Joan Mascaró and Marina Nespôr (eds.), *Grammar in Progress*, 447–454. Dordrecht: Foris.

430 *References*

- Vogel, Irene
1991 Level ordering in Italian Lexical Phonology? In: Pier Marco Bertinetto, Michael Kenstowicz and Michele Loporcaro (eds.), *Certamen Phonologicum II*, 81–101. Torino: Rosenberg & Sellier.
- Vogel, Irene
1993 Phonological interfaces in Italian. In: Michael L. Mazzola (ed.), *Issues and Theory in Romance Linguistics: Selected Papers from the Linguistic Symposium on Romance Languages XXIII*, 109–126. Washington, D.C.: Georgetown University Press.
- Vogel, Irene and István Kenesei
1987 The interface between phonology and other components of grammar: the case of Hungarian. *Phonology Yearbook* 4: 243–263.
- Vogel, Irene and István Kenesei
1990 Syntax and Semantics in Phonology. In: Sharon Inkelas and Draga Zec (eds.), *The Phonology-Syntax Connection*, 339–363. Chicago: University of Chicago Press.
- Watson, Keith
1997 French complement clitic sequences: a template approach. *French Language Studies* 7: 69–89.
- Wennerstrom, Ann
1993 Focus on the Prefix: Evidence for word-internal prosodic words. *Phonology* 10: 309–324.
- Wetzels, Leo
1992 Mid Vowel Neutralization in Brazilian Portuguese. *Cadernos de Estudos Lingüísticos* 23: 19–55.
- Wiese, Richard
1993 Prosodic phonology and its role in the processing of written language. In: G. Görz (ed.), *Konvens* 92, 139–148. Berlin: Springer-Verlag.
- Wiese, Richard
1996 *The Phonology of German*. Oxford: Clarendon Press.
- Williams, Edwin B.
1975 *Do Latim ao Português. Fonologia e Morfologia Históricas da Língua Portuguesa*. Rio de Janeiro: Tempo Brasileiro (3rd ed.). [Translation of *From Latin to Portuguese. Historical Phonology and Morphology of the Portuguese Language*. 1938. Oxford: Oxford University Press.]
- Zec, Draga
1988 Sonority Constraints on Prosodic Structure. Ph.D. Dissertation, Stanford University.
- Zec, Draga
1993 Rule domains and phonological change. In: S. Hargus and E. Kaisse (eds.), *Studies in Lexical Phonology*, 365–405. New York: Academic Press.

- Zec, Draga and Sharon Inkelas
 1991 The Place of Clitics in the Prosodic Hierarchy. In: Dawn Bates (ed.), *Proceedings of the Tenth West Coast Conference on Formal Linguistics*, 505–519. Stanford: CSLI.
- Zonneveld, Wim, Mieke Trommelen, Michael Jassen, Gösta Bruce and Kristján Árnason
 1999 Word-stress in West-Germanic and North-Germanic languages. In: Harry van der Hulst (ed.), *Word Prosodic Systems in the Languages of Europe*, 477–603. Berlin: Mouton de Gruyter.
- Zsiga, Elisabeth C.
 1992 A mismatch between morphological and prosodic domains: evidence from two Igbo rules. *Phonology* 9: 101–135.
- Zwicky, Arnold M.
 1977 *On Clitics*. Bloomington: Indiana University Press.
- Zwicky, Arnold M.
 1985. Clitics and particles. *Language* 61(2): 283–305.
- Zwicky, Arnold M.
 1987a Suppressing the Z's. *Journal of Linguistics* 23(1): 133–148.
- Zwicky, Arnold M.
 1987b French prepositions: no peecking. *Phonology Yearbook* 4: 211–227.
- Zwicky, Arnold M. and Geoffrey K. Pullum
 1983 Cliticization vs. inflection: English n't. *Language* 59(3): 502–513.
- Zwicky, Arnold M. and Geoffrey K. Pullum
 1985 Clitics and Particles. *Language* 61(2): 283–305.

Index

- Adjunction (prosodic), 19–20, 27–29
 of stressless prefixes, 166–170
 of proclitics, 196–203, 208–214
 Afar, 378n.185
 Allomorphy, 142, 145–146
 Ambisyllabicity, 34–37
 Arabic, 23
 Arbitrary gap, 128
 Autosegmental representations, 63
avos, 219–217

 Baule, 20, 23, 33, 262, 264, 331,
 382n.215
 Bengali, 7, 8, 23
 Biko, 152
 Brazilian Portuguese, 23, 44–45, 310,
 313–314, 328–328, 337–338,
 399n.336, 399n.337,
 405n.382, 405–406n.387
 Bulgarian, 146, 312–313, 402n.363

 Catalan, 23, 24, 358n.45
 Central vowel deletion (see Final
 central vowel deletion)
 Chamicuro, 23, 27
 Chi Mwi:ni, 406n.388
 Choctow, 23
 Chukchee, 133
 Class marker, 65
 Clitic climbing, 133
 Clitic cluster, 146–147
 Clitic group, 2, 17–19, 26, 31–32, 34
 Cliticization, 53, 58–59
 as phrasal affixation, 139
 Clitics (see also Pronominal
 clitics), 27–28, 139, 152, 173–
 175
 classification of, 173–182
 clitic clusters, 146–147
 diagnostics, 175–182
 disyllabic, 152, 175, 176, 177, 202–
 203, 309, 402n.358
 in the literature on EP, 42–43, 44–
 46, 50–51, 52
 incorporation (see Incorporation)
 list of EP clitics, 179
 prosodization (see Prosodization)
 (post)lexical insertion of, 127–128,
 326–327
 reduced/full forms of, 174, 180–
 182, 273–275
 reduction of, 273–324
 stressed, 384n.235
 syntactic (see Syntactic clitics)
 versus affixes, 204–208
 Clipping, 23–24, 250, 253–254, 337–
 338
 Compound Domain Hypothesis, 20,
 356n.27
 Compound intonational phrases, 20–
 21, 111, 113, 117, 375n.157,
 394n.301
 Compound prosodic words, 33, 218–
 249, 255–272, 262–272, 329–
 331, 340
 Compound words, 31, 32–33, 233–
 236, 236–240
 in the literature on EP, 43, 44
 prosodization of (see
 Prosodization)
 Coordination reduction (see Deletion
 under identity)

434 *Index*

- Deletion under identity, 22, 250–253, 254, 338
- Derived words
 in the literature on EP, 43, 46–47, 52–53
 prosodization of (see Prosodization)
- Direction of cliticization, 28–30, 41–42, 54–58, 184, 264, 313, 328–329
 with enclitics, 186–195
 with *mesoclitics*, 244–245
 with proclitics, 192–193, 195–203
- Dominance, 168
- Dutch, 20, 35, 160, 275, 276, 354n.5, 358n.43, 359n.49, 397–398n.324, 406n.388
 clitic hosts, 27, 402n.363
 clitics, 16, 18, 181, 274, 384n.236, 384n.237
 compound prosodic words, 261–262, 330–331, 397n.323
 Deletion under identity, 22–23
 direction of cliticization, 30, 328–329
 phonotactic generalizations, 22
 prosodic word, 25, 255, 256
 reduction, 276, 398–399n.334
 resyllabification, 24, 36, 359n.54
 word prosodization, 29, 32, 24, 34, 256
- Dyirbal, 175
- /e/-centralization, 78–82, 135
 with clitics, 205
 with prefixes, 167
 with suffixes, 205
- Emphatic stress, 72, 120–121, 357n.35
 in compound prosodic words, 227
 in prefixed words, 168
 with clitics, 201
- Enclisis (see also Pronominal clitics), 54–58, 154, 186–195
- English, 7, 8, 13, 15, 17, 165, 228, 355n.15, 357n.35, 382n.213
 clitics, 27, 181, 274, 277, 384n.236, 384n.237
 compound prosodic words, 261–262, 330–331, 340
 North American English, 7
 Old English, 402n.363
 phonotactic generalizations, 22, 383n.222, 383n.223
 prosodization of prefixes, 35
 reduction, 275–276, 398–399n.334, 405n.379
 resyllabification, 24
- Estonian, 23
- Extraprosodicity, 35
- Ewe, 12, 144
- Final central vowel deletion (see also Vowel deletion), 112
 with clitics, 189–191, 196–197
 within compound prosodic words, 226–227, 246
 within mesoclitic structures, 246
 within suffixed words, 226–227
- Final non-back vowel deletion (see also Vowel deletion), 61–62, 104–108, 160, 299–301
 with clitics, 186–188, 206–207
 with prefixes, 206–207
 with suffixes, 224–226
 within compound prosodic words, 224–227, 245, 247–249
 within mesoclitic structures, 245
- Final round vowel deletion (see also Vowel deletion), 108–112, 196–197
 with clitics, 188–189
 with suffixes, 226
 within compound prosodic words, 226–227, 226, 245–246

- Final round vowel deletion (cont'd.)
 within mesoclititic structures, 245–246
- Focal stress (see Focus)
- Focus, 3, 6, 120–121, 123–124, 357n.35
 on clitics, 179
 within abbreviations, 243
 within compound prosodic words, 222–223, 240, 243, 246–247
 within mesoclititic structures, 246–247
 within morphological compounds, 235–236
 within suffixed words, 222–223
 within syntactic compounds, 240
- Foot, 354n.11
 (in EP), 354n.11, 183, 336
- French, 12, 144, 146, 165, 228, 328, 355n.15, 358n.43, 406n.388
 Canadian French, 12
 clitics, 127, 311, 329, 384n.235
 compound prosodic words, 262, 331
 Middle French, 312
 Old French, 311–312, 313, 402n.363
 prosodic word, 255, 326, 357n.28
 prosodization of compounds, 33–34
 prosodization of prefixes, 392n.290
 word stress, 357n.28
- Frequency (word), 207, 249, 275, 303–309, 317–318, 389n.264
- Fricative voicing, 20
- Frisian, 312
- Galician, 355n.18
- German, 160, 354n.5, 359n.49, 406n.388
 clitics, 18, 27, 181, 273–275, 276, 312, 384n.236
 direction of cliticization, 30, 328–329
 deletion under identity, 22
 Minimal word requirements, 23, 35
 phonotactic generalizations, 22
 prosodization of clitics, 29
 prosodization of prefixes, 35
 prosodization of suffixes, 34, 256
 resyllabification, 24, 34
- Germanic languages 35–37, 181, 221, 265, 358n.44, 383n.222
 resyllabification, 24, 359n.54
- Glide insertion,
 obligatory, 83–85, 135, 104–205
 optional, 82
- Greek, 22, 24, 30, 194, 353n.4, 356n.27, 358n.43, 384n.235
 Ancient Greek, 175, 402n.363
 Demotic Greek, 17
 direction of cliticization, 30
 Modern Greek, 26, 144
 prosodic word, 25, 255, 326
 prosodization of compounds, 31–32
- Hausa, 12, 26, 144
 -haver, 149–150, 152–155
- Host
 phonological/prosodic, 18, 26–27, 184, 195, 264
 specialized, 127, 129–130
- Huave, 378n.185
- Hungarian, 6, 17, 25, 255, 326, 331
- Icelandic, 32, 358n.43
- Idiosyncrasy, 128, 129, 141–147
- Igbo, 384n.237
- Incorporation (prosodic), 77
 of enclitics, 186–195, 208–214
 of prefixes, 170–173
 of suffixes, 165
- Initial prominence
 prosodic word, 72, 118–119, 198–

436 *Index*

- Initial prominence (cont'd.)
 201, 318–319
 intonational phrase, 119, 309–314
- Initial vowels realization, 72–73, 92–99
 in suffixed words, 164
 in prefixed words, 167
 with clitics, 197–198
 within compound prosodic words, 227
- Interface
 lexical-postlexical level, 212–213, 263–264
 morphology-phonology, 263–263
 syntax-phonology, 4, 212–213, 263–264
- Interpolation, 132
- Intonational phonology, 376n.169
- Intonational phrase, 4–5
 initial position, 283–295, 197, 309–314
- Intonational phenomena (see also Tonal phenomena), 7, 8, 121–123, 123–124
- Irish, 23
- Italian, 7, 17, 18, 20, 35, 144, 228, 315, 355n.15, 355n.18
 clipping, 24
 clitics, 127, 146, 152, 311
 compound prosodic words, 330–331
 Minimal word requirements, 159
 prosodic word, 25–26, 255, 256, 326
 phonotactic generalizations, 22
 prosodization of clitics, 28, 328–329
 prosodization of compounds, 25, 32, 33, 358n.43
 prosodization of prefixes, 25, 26, 34, 35, 392n.290
 prosodization of suffixes, 256
 resyllabification, 36
 Standard Italian, 27, 28, 333
- Japanese, 23, 24, 340
- Juncture (see Sandhi)
- Kimatuumbi, 12, 144, 355n.19
- KiNande, 382n.215
- Kivunjo Chaga, 26
- Korean, 340, 359n.51, 384n.237
- Language acquisition, 339–340
- Latin, 17, 22, 24, 175, 326
 Vulgar Latin, 147, 401–402n.357
- Lardil, 23
- Letter names, 247–248
- Lexical component (see Lexical level)
- Lexical level, 10–17, 37, 59–60, 204–208, 332–335
 prosodization of clitics at, 182–183
- Lexical phonology, 10–17
- Lexical processes (see also Lexical level), 9–14
 in EP (summary), 124–126
- Linking vowel, 393n.293
- Lowering of stressless vowels, 85–89, 221
 /-velarization, 72
- Lucanian, 27, 28, 333
- Makonde, 356n.27
- Mapping (see Interface)
- Maximal
 intonational phrase, 20–21, 374n.157
 prosodic word, 225–226, 261
- Malayalam, 355n.15
- Mende, 12, 144
 -mente, 361n.65, 362n.71, 397n.319
 stress assignment, 43, 46, 48–49, 52–53
 prosodization, 219–227

- Mesoclis (see Mesoclitic structures)
 Mesoclitic structures, 130–131, 147–155
 direction of cliticization in, 244–245
 in the literature on EP, 49–53
 prosodization (see Prosodization)
 Metaphony, 371n.132
 Minimal
 intonational phrase, 20–21
 prosodic word, 225–226, 260, 325–326
 Minimal word requirements, 23, 158–159, 264, 382n.215, 396n.314
 Mismatch (structural)
 syntax-phonology, 8–9
 phonology-morphology, 15–16, 44
 Mobile affixes, 378n.185
 Monophthongization, 159

 Nasal final diphthongization, 74–78, 134
 with clitics, 206
 with prefixes, 206
 Nasal vowels, 64, 75
 Navajo, 139
 Neapolitan, 27, 28, 181, 193–194, 264, 333, 384n.235
 Nghanhacara, 29
 Ngiyambaa, 139
 Non-back vowel deletion (see Final non-back vowel deletion)
 Numerals, 247–249

 Organization of grammar, 10–17, 59–60, 332

 Pause insertion, 7, 8, 21
 Paradox, 15
 Particles (see also Clitics), 173
 Phonological phenomena, 6
 Phonological phrase
 in EP, 2–3, 5–6, 7, 337, 406n.389
 Phonological utterance, 4
 Phonological word (see Prosodic word)
 Phonotactic generalizations, 22, 158, 159–160, 329
 Pitch accent, 124, 161, 376n.169
 distribution, 3, 121–122,
 on clitics, 178, 194, 385n.241
 within compound prosodic words, 223–224
 within suffixed words, 223–224
 Polish, 15, 135–136, 380n.202
 Portmanteau forms, 138, 381n.207, 381n.212, 384–385n.239
 Portuguese, 358n.43
 Classical Portuguese, 311, 403n.366
 Medieval Portuguese, 311
 Postlexical level, 10–17, 37, 59–60, 204–208, 332–335
 prosodization of clitics at, 184–203
 Postlexical processes, 9–14, 332–334
 in EP (summary), 124–126
 Preboundary lengthening, 7, 8, 20–21
 Precompiled phonology, 12–13, 130, 145–147
 Primary stress (see also Word stress)
 in compounds, 31–32
 Proclisis (see also Pronominal clitics), 54–58, 192–193, 195–203
 triggers, 54, 131–134
 Prominence (see also Stress)
 assignment, 7
 compound prosodic word, 225–226, 235–236, 240, 243, 246–247, 260–262, 264
 intonational phrase initial, 309–314
 Pronominal clitics (see also Clitics), 54–57, 59
 affix/clitic ordering, 51, 137, 151
 and affixes, 58, 127–131

438 *Index*

- Pronominal clitics (cont'd.)
 distribution, 54–55, 127, 131–133, 153–155, 309–314
 lexical/postlexical attachment, 127–155
- Prosodic domain
 construction, 4–6, 10, 14–17
 strength, 335–337, 338, 406n.389
- Prosodic hierarchy, 2
- Prosodic phonology, 1–10
- Prosodic word
 boundary restructuring 36–37, 212
 construction, 24–26, 208–214, 255–272, 325–326, 405n.383
 diagnostics, 22–24, 157–163, 216–218, 336–337
 in the literature on EP, 60–62
- Prosodization
 of abbreviations, 241–243, 271–272
 of clitics, 18, 19, 25, 26–30, 186–203, 208, 195–203, 211–213
 of compounds, 24–25, 31–34, 134, 182–203, 263, 264, 326–329, 333
 of derived words, 24–26, 34–35, 163–173, 208–211, 219–227, 263, 228–232, 265–267
 of enclitics, 186–195, 208, 211–213
 of mesoclititic structures, 243–247, 270–271
 of morphological compounds, 233–236, 268
 of proclitics, 195–203, 211–213
 of stressless prefixes, 165–173, 208–211, 263
 of stressed prefixes, 228–232, 266–267
 of *stressed* suffixes, 219–227, 265–266
 of suffixes, 163–165, 208–211, 219–227, 265–266
 of syntactic compounds, 236–241, 269–270
 of words, 22–35, 208–214, 262–272
- Range, 124 (see also Tonal space)
- Reduction (see also Reduction phenomena)
 and allomorphy, 273–274, 277, 319–323
 and frequency, 275, 303–309, 389n.264
 and lexicalization, 319–323
 and word category, 290, 314–319
 of clitics, 283–298, 298–303
 hierarchy of (vowels), 280
 hierarchy of (categories), 290
- Reduction phenomena (see also Reduction), 177, 206–207, 273–281
- Register shift, 3
- Resyllabification, 16
 in Canadian French, 12
 in Dutch, 16
 in EP, 72, 80, 89, 91, 159–160, 334–335, 383n.228
 in Germanic languages, 24, 35–37
 in Romance languages, 24, 35–37, 334–335, 359n.54
- Rhythmic phenomena, 7
- Rhythmic stress (see also Secondary prominence), 200–202
- Romance languages, 35–37, 139, 140, 147, 221, 310, 328–329, 358n.44, 359n.49
 clitics, 173
- resyllabification, 24, 35–37, 334–335, 359n.54
- Romanian, 358n.43, 358n.45
- Round vowel deletion (see Final round vowel deletion)

- /r/-strengthening, 89–91
 - in suffixed words, 164
 - in prefixed words, 167, 168
- Sandhi, 8
- Sanskrit, 7, 22, 24, 255
- Schwa deletion, 72, 160, 278–280, 382n.221, 383n.227
- Sekani, 146
- Secondary prominence (see also Initial stress), 7, 118–119, 194, 199–200, 360n.59
- Segmental processes, 7
- Semivocalization, 99–103, 161
 - in prefixed words, 166, 228
 - V1, 85, 101, 102–103
 - V2, 84, 100–102
 - with clitics, 176, 186
- Serbo Croatian, 22, 26, 27, 175, 358n.45, 402n.363
- Shona, 382n.215
- Slavic languages, 358n.44
- Southern Calabrian, 384n.235
- Spanish, 24, 35, 36, 194, 255, 256, 311, 326, 328–329, 358n.43
- Speech rate, 276, 280, 283–291
- Speech style, 276, 400n.346
- Stem, 208–209, 211, 228, 391n.288
- Stress (see also Prominence)
 - echo (see Rhythmic stress)
 - intonational phrase level, 109, 120, 122
 - on function words, 256–257
 - on mesoclitic constructions (see also Word level stress), 258–259
 - on prefixes (see also Word level stress), 256–257, 259
 - on suffixes (see also Word level stress), 256–258
 - phonological phrase level, 109–110, 122, 406n.389
 - prosodic word level (see Word level stress)
- Stress clash, 3, 7, 8, 109–110, 112–113
- Stress window, 65, 88
- Strict Layer Hypothesis, 3, 397n.322
 - violations (/relaxation), 19–21, 27–28, 215, 331–332
- Structure preservation, 12, 14, 60
- Syllabification, 24, 35–37, 183, 369n.117, 370n.128
- Syllable
 - in EP, 183, 335–336
- Syllable degemination, 20, 114–118, 159
 - in prefixed words, 169, 392n.291
 - with clitics, 191–192, 196
- Syntactic clitics, 152–154
- Tagalog, 15, 152
- Theme, 208–210, 366n.97
- Theme vowel centralization, 69–70, 136
- Theme vowel deletion, 73–74, 136–137, 150
- Tonal phenomena (see also Intonational phenomena)
 - tonal space (see also Range), 121
 - tonal association, 121–123
 - lexical, 123
- Transparent/opaque structures, 35, 165–166, 170–173, 228–229, 267–268, 269, 394n.299
- Truncation (see Clipping)
- Tune association, 7, 122
- Turkish, 17, 22, 24, 194, 359n.51
- Tzotzil, 26
- Underlying system
 - of EP, 63–64
- Underspecified segments, 75, 80, 89, 94–97, 366–367n.99

440 *Index*

- Vowel deletion
 - central (see Final central vowel deletion)
 - non-back (see Final non-back vowel deletion)
 - round (see Final round vowel deletion)
 - schwa (see Schwa deletion)
 - with clitics, 176–177
 - with prefixes, 161
 - with suffixes, 164
- Vowel harmony, 368n.107
- Vowel reduction (see also Reduction),
 - 67–73, 97, 161, 278
 - lack of/exceptions to, 47, 68–71, 85–89, 121
 - with clitics, 175–176
 - with prefixes, 228
- z-evaluative suffix, 43, 46–47, 48–49, 219–227, 361n.65, 362n.71, 380n.201
- Word (morphological), 208
- Word level stress (see also Primary stress), 22, 48–53, 121, 150–151, 161, 255–259, 357n.28
 - and the prosodic word, 18, 260–262
 - assignment, 48–49, 52, 64–67
 - on clitics (lack of), 175–176, 193–194
 - within mesoclitic structures (see also Mesoclitic structures), 243–244
 - within suffixed words, 164, 219–221
 - within prefixed words, 166, 170, 228
- Weight, 3, 5
- Yidin^y, 23