# The Influence of Phonology on Inflection

## The interplay between syllabification and lexical insertion in Pallarese Catalan

Eduard Artés Cuenca

PhD Thesis

Supervisor: Dr. Eulàlia Bonet Alsina

*A la meva mare*

Gender is a complexity whose totality is permanently deferred,
never fully what it is at any given juncture in time.

*Gender Trouble*
Judith Butler

# ABSTRACT

This dissertation deals with the constraints that phonology imposes on the exponence of gender. In particular, it explores the interaction between epenthesis and nominal inflection in Pallarese Catalan clitics. Since inflection is located word-finally in Pallarese, I argue that right-edge epenthesis is avoided in order to maintain a strict correspondence between morphological and phonological structure. Inflectional vowels may be used for syllabification purposes in word-final position instead.

It is assumed in this thesis that every syntactic functional head projects a post-syntactic theme position (Oltra-Massuet, 1999) where gender is realized. A key aspect of the proposal is the idea that the exponents of gender are floating features, and thus inflection proceeds in two steps. First, the theme position is spelled out with an underspecified vowel (i.e., Th ↔ V) and, subsequently, the floating features associated with gender ([+fem] ↔ [+low], i.e., -[a]; [–fem] ↔ [+labial,–high], i.e., -[o]) are attached to the V-slot of the theme position. Couched within Optimality Theory, I argue that the constraint ranking determines whether the featural gender exponents surface or not. The feminine is always realized due to a constraint that favors parsing [+low] (which implies adding association lines that are absent in the input), whereas in the masculine the insertion of new association lines in the output is dispreferred and the features associated with [–fem] are not parsed, which accounts for default masculine -[Ø] exponence. If gender is part of the morphosyntatic composition of a clitic, as in the 3rd person singular masculine accusative clitic, these floating features can nevertheless be attached to the V-slot of the theme position to improve syllabic structure under certain phonotactic conditions. This 'morphological solution' is less costly than (regular) word-initial epenthesis because it does not need to create a new skeletal position or insert new features.

Impoverishment (Bonet, 1991) deletes gender features in the 3rd person plural accusative clitic, and thus the corresponding phonological features associated with gender cannot be used for syllabification purposes. The theme position with the

V-slot is maintained, though, and the default epenthetic features of Pallarese are inserted when required by phonotactics, which forces a thematic interpretation of this vowel. This solution is preferred over word-initial epenthesis because the theme position already provides a skeletal slot. The same procedure applies to other clitics that do not bear gender features either.

Even though Pallarese shows a complex morphophonological intertwining regarding gender exponence, the OT analysis presented in this dissertation makes exclusive reference to phonological objects. The morphosyntactic structure of the nominal system constrains epenthesis, but strict modularity can be maintained. Furthermore, the use of floating features in the input can dispense with gender allomorphy (cf. Bonet et al. 2007).

As for nouns and adjectives, the general process that spells out an underspecified vowel in the theme position, on the one hand, and floating place features for gender values, on the other, is only valid for default endings. That is, vowels other than *-a* (feminine) and *-o* (masculine) —when it surfaces— cannot be considered gender markers and need to be fully specified in the theme position of lexical entries instead. Therefore, nouns with non-regular endings are stored as complex representations and phonologically realized in one single step. This supports theories that assume that one exponent (or set of exponents) can spell out whole morphosyntactic structures (e.g., Siddiqi 2009, Caha 2009 or Bermúdez-Otero 2012).

# ACKNOWLEDGEMENTS

The time to thank all the people that have helped me in one way or another during this long journey has (finally!) come. Words cannot express how grateful I am to Eulàlia Bonet. She is the best supervisor one can ask for, and this thesis would not have been possible without her. I want to thank Eulàlia for her patience and availability in reading all the chapters of this thesis and answering all my questions (always with a smile), even the dumbest ones. Besides being an awesome linguist, Eulàlia is a humble person, something I truly appreciate.

I am also very grateful to the members of my thesis committee. My heartfelt thanks to Joan Mascaró (who has always been willing to help during my life as a PhD student at UAB) and Maria Rosa Lloret. I am particularly indebted to "Xico" Torres-Tamarit. Xico always encouraged me to submit abstracts to conferences to overcome my insecurities and, on many occasions, he had more faith in me than I did myself. In the end he was right and I was able to finish this thesis!

I would like to thank all the people at the *Centre de Lingüística Teòrica*, where I felt at home during the four years of my grant, especially Mar Massanell (for her support with the fieldwork) and Teresa Cabré and Maria Ohannesian (both members of the Phonology Reading Group).

I could not have survived the PhD without the countless coffee and lunch breaks at UAB with my colleagues: Teresa Blasco, Isabel Castro, Ekaterina Chernova, Adriana Fasanella, Javi Fernández, Jordi Fortuny, Elías Gallardo, Lucía Medea García, Yurena Gutiérrez, Maya Leela, Ares Llop (who also gave me the topic for this disserttion!), Cristina Real, Marina Roman, Carlos Rubio, Io Salmons and Qiuyue Zhong. I really miss my UAB life and the great moments we all spent together.

I spent the Fall semester 2013 at the incredible MIT Department of Linguistics. I want to offer my thanks to Donca Steriade, who welcomed me into the department and transmitted her passion for phonology to me, and to Maria Luisa Freitas and Elena Nulvesu, for making a new place feel more familiar.

# TABLE OF CONTENTS

# ONE

# INTRODUCTION

## 1.1 The phonology-morphology interface

This dissertation is about the phonology-morphology interface. More specifically, it is concerned with the mechanisms that control gender exponence and how they are related to syllable well-formedness requirements. To correctly understand this set of interactions we first need to determine how the overall architecture of grammar works. Morphology and phonology need to be clearly defined in what their duties are, as well as their relation to syntax and the lexicon. In this section I will provide an outline of the advances that have been made in this area of Generative Linguisticss over the past decades, whilst justifying the assumptions that I make for the Catalan data that are analyzed in the chapters that follow.

### 1.1.1 Morphology and the lexicon

Classic generative approaches assume an architecture of grammar as in (1):

(1)

*Lexicon*

$\downarrow$

*Syntax*

*PhonologicalForm*          *LogicalForm*
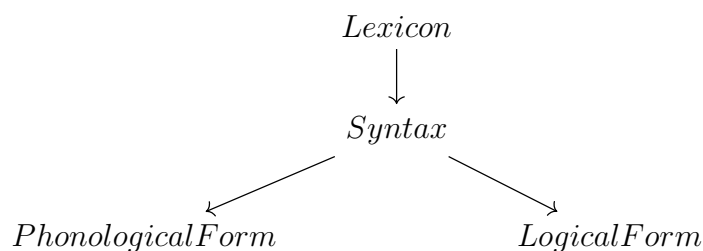
Words are associations of sound and meaning, but how these associations are stored and what their structure is is a matter of debate for every linguistic theory. In fact, one of the first problems that we face when confronted with a schema like

(1) is to define what the lexicon is. Generativism assumes that the output of syntax is interpreted by the Conceptual-Intentional and Articulatory-Perceptual systems in its Logical and Phonological Forms, respectively, thus defining and delimiting what is done at both the semantic and the phonological modules. However, what the input to syntax is, and which role we should assign to morphology prior to the building of syntactic (phrase) structure, is more controversial.

As Fábregas & Scalise (2012) point out, the use of the term 'lexicon' in the literature can lead to some misunderstandings, and they separate this concept from that of morphology as follows:

> Morphology is the set of operations that manipulate morphological units in order to form new words or adapt the form of words to specific syntactic contexts. (...) The lexicon, in contrast, is a list of units which is stored somewhere in the mind of the speaker; these are the units combined in syntax (and/or morphology). The lexicon, in a strict sense, does not have generative power and does not create new objects. Its role is to store the units that morphology and syntax can manipulate in their operations. (Fábregas & Scalise, 2012, p. 5)

This distinction is particularly useful for Lexicalist approaches to language, which sometimes conflate these two concepts. Lexicalism considers building words and building phrases to be essentially different operations. And it is precisely in the lexicon that these distinctions arise, as some morphological operations take place there. That is, at least for derivation, words are assembled in the lexicon (assuming this Lexicalist 'mixed' view of the lexicon) and they are later sent to syntax, where phrase structure is built. Therefore, taking as an example Catalan deverbal nouns, it is in the lexicon that a word like *assumpció* 'assumption' (derived from the verb *assumir* 'to assume') is created, and it is also in the lexicon that the verb *assumir* is specified to take *-ció* (and not *-ença*, e.g.) as a suffix to become a noun.

Later developments in the generative tradition, however, have redefined the relation between syntax and the lexicon, and, as a consequence, the way morphology and phonology interact. One of the most developed and influential non-lexicalist theoretical frameworks is Distributed Morphology (henceforth DM; Halle & Marantz 1993), which actually denies the existence of what has traditionally been considered morphology altogether. For DM, all morphological operations are in fact carried out in the syntactic component, and there is, thus, no distinction between building words and building sentences. The architecture of grammar in DM is as in (2) (reproduced from Harley & Noyer 1999, p. 3):

(2)
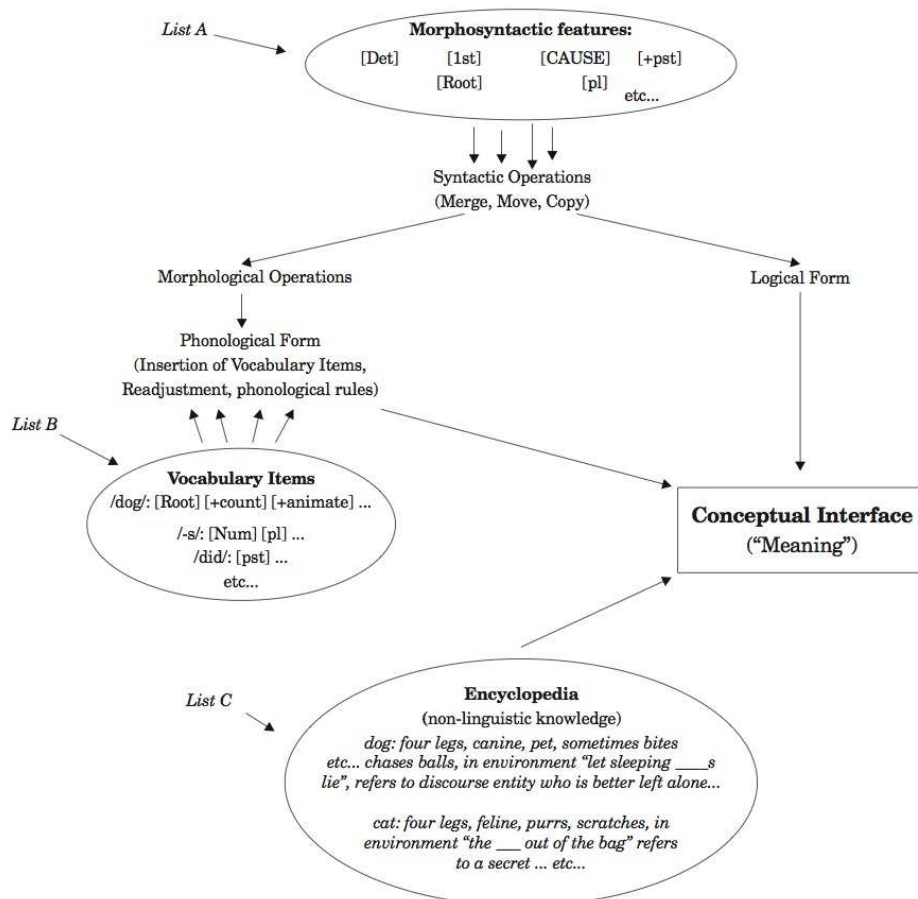


In (2) it is shown that there is no proper lexicon in DM. It is instead 'distributed' (hence the name) in different lists; only List A and List B matter for the present purposes (list C contains encyclopedic knowledge). As Bobaljik (2012) points out, there are two leading ideas in DM, which are reflected by lists A and B:

1. Syntax-all-the-way-down

2. Late Insertion

List A consists only of morphosyntactic features (without any phonological content) which undergo syntactic operations both below and above the word level; thus the idea of syntax-all-the-way-down, as there is no difference between morphology and syntax. These discrete units that belong to List A, though, will only get phonological content in the mapping to Phonological Form (i.e., the Spell-Out process), and hence the idea of Late Insertion, because phonological content is only provided at the end of the syntactic derivation. This is in fact what List B contains, the so-called Vocabulary Items, which would be the closest notion to lexical items of

Lexicalist approaches, although essentially different: they are only *instructions for insertion*. That is, Vocabulary Items give phonological content to the morphosyntactic features that have been merged through the syntactic derivation, but they are, crucially, not stored with these features. An example can be seen in (3):

(3)     [+fem] ↔ /a/

In Romance languages like Catalan, Italian or Spanish, feminine gender is spelled out as *-a*, e.g., *casa* 'house'. However, according to DM, the Vocabulary Item in (3) is not stored in the lexicon as a whole, rather it just provides the association of /a/ to [+fem], which allows its licensing. The insertion of /i/ (assuming it is also a Vocabulary Item), instead, is not licensed because it is not appropiate in the context [+fem], but in principle nothing goes against it (it competes with /a/). Thus, the difference with regard to Lexicalism is substantial.

Even if word-formation is syntactic, though, there is still something that is reminiscent of the traditional morphological module in DM (although very limited in its functions), which appears under the label 'morphological operations' in (2). These are the set of operations that take place post-syntactically and include Fusion, Fission and Impoverishment, among others. Only Impoverishment (Bonet, 1991) is treated in this dissertation.

Impoverishment is the process by which syntactic features are deleted after syntactic derivations, i.e., at the spell-out of syntax but prior to Vocabulary Insertion. As a consequence, the insertion of exponents for a given morphosyntactic feature or set of features that have been impoverished is not licensed, as the context provided in Vocabulary Items is no longer met (the exponents associated with the deleted syntactic features cannot be phonologically spelled out). In certain languages there are gender distinctions in the singular but not in the plural. Under an Impoverishment analysis, a rule like (4) explains these asymmetries:

(4)     [±fem]⟶ Ø /__[+pl]

The behavior of a language, for example, with two gender distinctions in the singular, masculine and feminine, but no differences in the plural, is easily captured by the rule in (4). Any plural noun will be constructed syntactically and given a gender value. At the spell-out of syntax, however, the [±fem] feature will be deleted and, therefore, no exponent will be inserted. This implies that gender is neutralized in the plural, as no difference between masculine and feminine will arise. This idea is crucial for this dissertation, and in fact something similar to (4) will be proposed for the gender exponence of Pallarese Catalan 3rd person plural pronominal clitics

in 5.1.2.

The above assumptions made by DM will be employed throughout this thesis. In particular, the notion of Late Insertion is crucial for the analysis, as well as the Impoverishment operation that takes place before lexical insertion. However, it must be clearly stated that this does not imply that all the tenets of the theory are held. The focus of the study is the Catalan nominal system (more specifically, the clitic system of Pallarese, analyzed in chapters 5 and 6), and there is nothing to be said about the assumptions made for phrasal syntax and semantics, or about the moves taken by every specific version of DM. What matters, as already mentioned, are the concepts of Late Insertion and Impoverishment, and the analysis should hold in any framework that uses these notions. Furthermore, the phonological analysis is implemented in OT, which makes very different assumptions about the overall architecture of grammar. Both theories, though, can be maintained in their respective domains, as there is no incompatibilty with assuming a serial model of morphosyntax and a parallel model of phonology. In fact, they have been used together in several analyses (cf. Trommer 2002; Svenonius 2012). Morphosyntax and phonology have different properties, and thus different theories may be suitable for each language module.[1]

Given the general properties just sketched, DM can be grouped within lexical-realizational frameworks of morphology (in Stump 2001's terms), a framework also defended in this work. Some theoretical proposals made in the literature for inflection and the reasons for the choice of a lexical-realizational model over other options are discussed in the next section.

### 1.1.2 Theories of Inflection

Inflection has been a privileged area of study in linguistics due to its implications for the overall architecture of grammar. Hockett (1954) first introduced the terms Item-and-Arrangement (IA) and Item-and-Process (IP), which have been popular to describe the 'two models of grammatical description' (as he establishes in the title of his article) ever since.[2]

---

[1] "(...) parallel computation within a module does not preclude, or make any statement, regarding the eventually serial communication among computational systems." (Scheer, 2012, p. 67)

[2] Hockett acknowledges that there is a third option, the Word-and-Paradigm model "which is older and more respectable than either <IA or IP>, (...) the traditional framework for Latin, Greek, Sanskrit, and a good many more modern familiar languages."(Hockett, 1954, p. 386). Word-and-Paradigm models consider the word to be the minimal unit of meaning. That is, they can use the notion of morpheme for descriptive purposes and refer to sub-units of the word in analytic terms, but do not consider these sub-units to be meaningful entities. Only at the word

The main difference between these two systems is, in general terms, the status they give to the notion of morpheme. Morphemes are considered 'real' entities in IA models, whereas IP models only use them as descriptive tools, because changes in inflectional paradigms are seen as the result of rule application to a base form. In other words, "the essence of IA is to talk simply of things and the arrangements in which those things occur. One assumes that any utterance in a given language consists wholly of a certain number of minimum grammatically relevant elements, called morphemes, in a certain arrangement relative to each other." (Hockett, 1954, p. 387). By contrast "<in IP> a derived form consists of one or more underlying forms to which a process has been applied" (p. 396). The use of the terms 'arrangement' and 'process' in both theories seems clear, as it captures the different structural assumptions on word formation, which are shown in (5) and (6) for the plural in English (examples taken from Bonet 2008a, p. 1):

(5)

root                            [+plural]
 |                                  |
 |                                  |
/dɔg/                             /z/
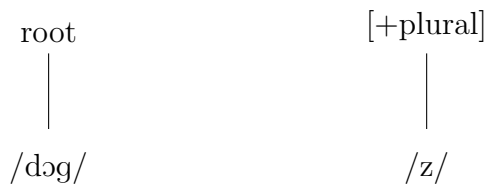
(6)     a.    Plural = /X/ > /Xz/
        b.    /dɔg/ > /dɔgz/

Whereas in (5) the suffix /z/ is an autonomous element which corresponds to [+plural], in (6) the whole rule is the notion 'plural'. That is, the scheme /X/ > /Xz/ is stored as a rule which converts singular nouns into plurals, but crucially /z/ on its own does not convey plurality.

Both approaches are better suited to explaining certain inflectional processes rather than others, and while concatenative exponence seems more straightforward in IA models (like DM), root-and-pattern morphology (common in Semitic languages like Arabic), by contrast, can be more easily accounted for using an IP framework (cf. Anderson 1992; Aronoff 1994). No definitive argument exists that allows us to reject one of the theoretical approaches in favor of the other, but an IA model is

---

level a suffix has some value, never on its own, and it is precisely by relating elements of similar form to one another that they get significance —hence the name Word and Paradigm— (Anderson, 2015). Under this analysis, in a Catalan word like *perillós* 'dangerous', decomposable into *perill* 'danger' plus the suffix *-ós*, this last element does not have any value in its isolated form, as it only gets real significance when attached to a base. I do not aim to provide a thorough analysis of all inflectional models of morphology, and therefore Word-and-Paradigm frameworks will be left out of the discussion from now on.

assumed here because it can deal with the behavior of masculine gender exponence in the Catalan nominal system, the main focus of this thesis.

As pointed out by Bonet (2008a), process-based accounts cannot easily explain the behavior of a set of masculine Catalan plurals whose base ends in a sibilant, like *gossos* 'dogs':[3]

(7)
| | | | | |
|---|---|---|---|---|
| *cor* | 'heart' | > | *cors* | 'heart' |
| *gos* | 'dog' | > | *goss**os*** | 'dogs' ([gósus]) |

In Catalan, plurals are created by the adjunction of -*s*. As shown in (7), in masculine default nouns like *cor* (with a -Ø inflectional ending), this yields *cors*, as expected. By contrast, in singular nouns like *gos*, whose root ends in a sibilant, an unexpected vowel, -[u], appears in the plural. This is in fact the vowel that appears in some marked nouns which do not follow the masculine default pattern (i.e., -Ø inflection), such as *moss*[u] 'guy', but in *gossos*-like plurals is used as a means to repair syllable structure (\**goss*, as a sequence of two consecutive sibilants as codas is not tolerated in Catalan). Therefore, we are forced to conclude that this vowel is epenthetic in nature, although it must be morphologically conditioned because otherwise the default epenthetic vowel of Catalan, schwa, would surface in this context, \**goss*[ə]*s*.[4]

The argument provided by Bonet (2008a) is the following. In IP models, word-formation rules have to apply sequentally (in a strict order) at the output of syntax, as words are just bundles of unordered features attached to a root. In the plural cases that occupy us here, this means that two different rules are needed. First, a rule like (8) that spells out inflectional class is needed:

(8)     *cor* [kór] vs. *mosso* [mósu]

a.
$$/X/ \quad > \quad /X/$$
$$/kɔr/ \quad > \quad /kɔr/$$

b.
$$/X/ \quad > \quad /Xu/$$
$$/mos/ \quad > \quad /mosu/$$

For class I, the rule states that the root and the stem remain the same, but for

---

[3]These cases will be treated in depth throughout this disseration, as they are a puzzle in all analyses of Catalan nominal morphology (in this section they are only presented as a means to justify the choice of one inflectional model over the other). See 3.1.1 for the analysis of theses cases in the literature, 4 for the sketch of the proposal presented in this dissertation which accounts for nouns like *gossos* and 7.1 for the specific morphophonological analyisis of this set of nouns and adjectives.

[4]The word *goss*[ə]*s* does actually exist in Catalan, but it is used to refer to female dogs (built upon the feminine base *goss*[ə] 'female dog', which bears the feminine marker -*a*). In the masculine —or neuter— plural only *gossos* is possible.

class II, instead, it adds -[u] to the root, thus creating a distinction between *cor* and *mosso*.

The second formation rule regards plural, and it is like the one given in (6), also valid for Catalan. By adding this second rule after the inflection class I, we predict the correct derivation of *cors* and *mossos*, (9), but not of *gossos*, (10):

(9)
$$/kɔr/ \quad > \quad /kɔr/ \quad > \quad /kɔrs/$$
$$/mos/ \quad > \quad /mosu/ \quad > \quad /mosus/$$

(10) $\quad /gos/ \quad > \quad /gos/ \quad > \quad */goss/$

The problem with the plural *gossos* is that the noun *gos* belongs to class I as it does not get the -[u] ending in the singular (*\*goss*[u]), as opposed to *moss*[u], which belongs to class II and gets the ending added to the root $\sqrt{\text{MOS}}$). Without this intermediate step, the plural formation rule can only attach *-s* to the stem, which remains *gos* after the realization of class, thus yielding *\*goss*. The problem originated by the sibilant contact appears after the realization of class, and, crucially, the rule that realizes class cannot apply again.

As pointed out by Bonet (2008a, p. 7), the solution could be left to the phonology, but then we need to explain the vowel quality in the epenthetic element. That is, if this vowel is considered to be purely epenthetic, the spell-out rules for inflectional class and plural can be maintained in their correct order, but the problem still remains as the appearance of this vowel is morphologically conditioned. As previously mentioned, the default epenthetic vowel of Catalan is schwa, and the use of -[u] in this context cannot be justified on phonological grounds.

If the marker -[u] is considered, instead, an independent morpheme —as it would be in IA models— it can be used as a repair strategy at different stages of the morphological derivation, as we are not constrained by rule-ordering. This approach will prove in fact to be the most successful in the following chapters, where it is shown that the phonological features associated with [–fem] remain floating at the phonological component —after Vocabulary Insertion— and they are only linked to the skeleton and sent to phonetic interpretation under certain phonotactic requirements, like those provided by the contact of two sibilants in *gossos* (*\*goss* > *gossos*).

The traditional distinction between Item-and-Arrangement and Item-and-Process theories is, in some ways, refined in Stump (2001), who proposes the following taxonomy for inflectional theories based on two axes:

- lexical vs. inferential

- realizational vs. incremental

On the one hand, inflectional theories can be either lexical or inferential. Lexical theories of inflection assume that the associations between an element's syntactic properties and its morphological correlate are stored in the lexicon. That is, *-a*, the default feminine inflectional marker in Catalan, is paired in a lexical entry with the specifications 'feminine, singular'. For inferential theories, by contrast, there are no such things as morphemes (i.e., inflectional markers do not exist as separate objects); instead, morphological rules incorporate inflectional exponents to a word's root. Therefore, lexical theories correspond to IA models and inferential theories to IP models.

On the other hand, theories can be either realizational or incremental. In incremental theories, inflection is 'information-increasing', which means that inflectional markers are responsible for adding certain morpho-syntactic features to a stem. Crucially, these features would otherwise be absent, as opposed to realizational theories, for which all the morpho-syntactic information is already present in the stem and it only needs to be fulfilled phonologically. Thus, in incremental theories, a word like *casa*, which bears the default feminine ending *-a*, only gets the values {[+fem],[–pl]} when this suffix is attached to the root. In realizational theories, instead, the word already carries these morphosyntactic features, i.e., *cas*.FEM.SG., but the phonological features associated with it are only inserted later.

Given these distinctions, and as previously mentioned, the model which is assumed throughout this dissertation is lexical-realizational. Lexical because the morphosyntactic features stored in the lexicon constitute 'abstract morphemes', and realizational because these features are subsequently provided phonological content after syntactic derivation (i.e., the late insertion idea). Besides the fact that a late insertion model better captures the behavior of Pallarese Catalan, it is worth mentioning that several psycholinguistic studies have demonstrated that phonological content is accessed separately from morphosyntactic information. The results provided by Caramazza & Miozzo (1997) and Vigliocco et al. (1997) show that unimpaired Italian speakers can retrieve the gender of a noun but not its phonological properties in a tip-of-the-tongue state, and the same results have been found for anomic patients (Badecker et al., 1995). If both morphosyntactic and phonological information were stored together, speakers should not present these asymmetries in lexical retrieval, and thus these data constitute strong evidence for lexical computation in separate stages.

Until now, we have discussed different proposals on how morphology (or syntax) builds inflected words. This dissertation, however, is mainly focussed on inflectional

*exponence*, which "may be understood declaratively as a systematic phonological relation between inflected word forms sharing a set of inflectional features and their bases" (Trommer & Zimmermann, 2015, p. 47). It makes certain (crucial) assumptions on the morphological structure of the nominal system, but it mostly deals with the realization of gender, i.e., its phonological manifestation. According to Trommer & Zimmermann (2015), there are three different kinds of inflectional exponence:

1. Additive exponence

2. Transformational exponence

3. Templatic exponence

Additive exponence is the most well-studied kind of exponence, which consists in adding a piece of phonological material to a stem, for example joining -*s* to a nominal category in Catalan plural formation, as we have previously seen, e.g., *cor* 'heart' > *cors* 'hearts'. Transformational and templatic exponence are traditionally seen, by contrast, as not always concatenative. As indicated by the term, transformational exponence is indicated by the phonological transformation of a stem, e.g., imperative in Hausa, where the tonal contour of the verb is overwritten by the melody Low-High: *kwá:ná* > *kwà:ná* 'spend the night'; *táʃ-ì* > *tàʃí* 'get up'. Templatic exponence, instead, is expressed by imposing a fixed phonological pattern onto a stem, as in the plural of some German nouns that have to adapt to a disyllabic foot, e.g., *Kind* > *Kinder* 'child, children' vs. *Mieder* > *Mieder* 'bodice, bodices' (examples taken from Trommer & Zimmermann 2015, p. 48).

Catalan nominal inflection, which is expressed by inflectional markers at the right edge of the word, belongs to additive exponence. It is commonly analyzed as a segmental process because it consists of the concatenation of morphs. However, Autosegmental Phonology (Goldsmith, 1976) allows an analysis of the Catalan nominal system that does not assume the storage of full segments as exponents of gender. In chapters 4 and 5 it will be shown how Catalan gender exponence can, in fact, be analyzed as featural affixation (Akinlabi, 1996). By completely dissociating phonological realization from a purely morphological inflectional position (named 'theme position' in this dissertation after Oltra-Massuet 1999) we can consider floating features to be the actual exponents of morphosyntactic gender values. These features will only be linked to the skeletal position provided by the theme position —which contains only an underspecified vowel with no place features— under certain conditions and restrictions imposed by the OT constraint ranking. With this move, some inflectional phenomena can be analyzed as purely phonological. That is the case, for

example, of the alternation between -Ø and -o in singular-plural pairs like *gos-gossos* 'dog-dogs' mentioned above. While most analyses (see section 3.1.1) assume that there must be some kind of lexical specification or allomorphy in this set of nouns to account for the alternation, the use of floating features can derive these nouns from the general procedure of gender exponence. Syllable well-formedness requirements determine the linking of -o (i.e., the featural exponents that correspond to -o) to the skeleton to avoid the contact of two sibilants (\**goss*), but neither additional changes in the general pattern of exponence nor allomorphy are needed.

Exponence is, thus, the relation between morphosyntactic and phonological features. From a modular point of view, this implies that the information from one module must be translated into another. How this process is carried out is the topic of the next section, 1.1.3, where the duties of phonology and morphology are better defined according to the assumptions already made for the chosen theoretical model of inflection.

## 1.1.3 Morphology, phonology or both?

Within a lexical-realizational model of inflection such as the one adopted here, the lexicon, in a strict sense, consists of bundles of morphosyntactic features, and only after syntactic derivation are these features given phonological content. We have a clear separation between different modules of the grammar (Fodor, 1983), but the question that needs to be addressed is *how* these two modules —morphosyntax and phonology— talk to each other, as Scheer (2012, p. 13) points out:

> (...) are morpho-syntax and phonology distinct computational systems whose input are distinct sets of vocabulary items? If so, in order to be able to communicate at all, the output of morpho-syntactic computation needs to be translated into phonological vocabulary before phonological computation can proceed.

It is clear that morphosyntax and phonology speak different 'languages', and thus the features interpretable by the syntax need to be converted into phonological material 'before phonological computation can proceed'. This distinction between the kinds of computations performed in every module is crucial. On the one hand, "syntax provides structures that can be linearized and translated into different kinds of constituents" (Bye & Svenonius, 2012, p. 431). On the other hand, phonology operates only with phonological objects (features, syllables, association lines,...), which means that, besides prosody —which interacts directly with syntax— there is no access to morphosyntactic information by phonology. Alignment constraints

can refer to edges of both morphosyntactic and prosodic structure, but, other than that, no prior information of the syntactic computation is sent to the phonological component. What is left, then, to the traditional morphological component? As already seen, some postsyntactic operations like Impoverishment take place in the remains of the morphological component but, most importantly, morphology is responsible for the connection of morphosyntactic values with phonological exponents. It can 'read' both languages but, crucially, it cannot operate on phonological representations. This is what (Bermúdez-Otero, 2012, p. 44) labels 'Morph Integrity Hypothesis':

> The representational currency of morphology is the morph: morphology is not allowed to operate directly upon elements of phonological representation such as features, segments, nodes, or association lines.

This is the idea pursued in this thesis, which adopts a conservative view of phonology and maximally limits the interaction with morphology. Morphology's attribution is, essentially, the concatenation of morphs.[5] A consequence of the Morph Integrity Hypothesis, as pointed out by Bermúdez-Otero (2012, p. 53), is that all morphology is concatenative, an idea adopted in other recent proposals (Bye & Svenonius, 2012; Trommer & Zimmermann, 2015) —and first introduced by Lieber (1992)— which also holds this view of morphology as a transfer module that connects syntactic structure to phonological content, whatever this may be (segments or units bigger or smaller than segments). It is not the intention of this thesis to prove that all non-concatenative phenomena (e.g., subtraction, ablaut, ...) can actually be analyzed as cumulative exponence, but it is worth noting that the data analyzed here fit perfectly into the proposal. The realization of gender in Catalan is analyzed (counter intuitively) in a non-linear way which can be extended to other non-segmental processes in different languages, as the exponents of masculine and feminine gender values consist of floating features (see chapter 4). Therefore, the proposal is useful for both segmental and non-segmental phenomena and thus compatible with the Generalized non-linear affixation model (Trommer, 2011; Bermúdez-Otero, 2012).

Given this assumption concerning morphology and having reduced its scope, it is clear that most things attributable to morphology must take place in other language components. In fact, all the constraints argued for in this thesis make reference exclusively to phonological objects like features, association lines, etc., and some of the alternations found in Catalan inflection can be accounted for only by

---

[5]This is a *desideratum*, because there are certain phenomena, like clitic combinations, that are not easily explained without a more richly endowed morphology. As this is not a thesis about the architecture of grammar, I will not concentrate on this topic.

phonological means. The analysis of Pallarese Catalan gender exponence assumes, as already mentioned, that floating features are the exponents of [±fem] values, and two kinds of OT constraints will be in charge of generating the expected outputs:[6]

- PARSE-FEATURE

- DEP-LINK (Morén, 2001)

These constraints can account for, among other things, the alternation found in *gos-gossos* 'dog-dogs'. It will be shown in 5.1.1 how these 'conflicting' constraints (together with other phonotactic constraints), which work in opposite directions, explain the presence of *-o* in the plural contexts. Floating features are part of the input for inflection, but most of the time these features will remain floating in the masculine because DEP-LINK, which militates against the insertion of association lines that are not present in the input, is ranked higher than PARSE-FEATURE, which favors the phonetic interpretation of the floating features (i.e., it favors candidates with phonological features connected to skeletal positions via association lines). Only under certain phonotactic conditions, like those provided by the contact of a sibilant belonging to the root and the plural morph *-s*, will the linking of floating features be licensed and preferred over epenthesis.

This is just a simplified sketch of the analysis presented in 5.1.1, but what should be emphasized is that there is no need to resort to a complex intertwining of phonology and morphology as developed in recent proposals of Catalan nominal inflection like Bonet et al. (2007). By assuming a nonlinear approach to gender exponence, the output of morphology is always the same under the present analysis, and it does not need to include allomorphs or morphological constraints in the OT phonological ranking, as Bonet et al. (2007) propose. The leading idea in Bonet et al. (2007) is that the lexical entry for gender consists of a multiple input of hierarchically ordered allomorphs (e.g., /{Ø>u>ə}/ for the masculine), and it is in the phonological module that one of these allomorphs is finally chosen over the others. Thus, it is the OT ranking which finally determines the inflectional marker due to the action of the constraint PRIORITY that gives preference to the highest-ranked allomorph of the hierarchy (see 3.1.1 for a detailed review of the proposal). Some phonological constraints dominate PRIORITY, though, and therefore candidates that bear the highest-ranked allomorph are not always the winners. This explains the asymmetry

---

[6]A standard parallel version of Optimality Theory is assumed throughout this thesis. Cyclic versions of OT like Stratal OT (Bermúdez-Otero, 1999; Kiparsky, 2000), inspired by the ideas of Lexical Phonology and Morphology (Kiparsky, 1982b), could also account for the Catalan data analyzed here, but it is beyond the scope of this dissertation to discuss the pros and cons of serial and parallel models of phonology.

between *gos*[Ø] and *goss*[u]*s*, as in the plural context the phonotactic restriction against two consecutive sibilants forces the insertion of -[u], ranked second in the hierarchy. The unmarked allomorph -[Ø] can only appear in the singular, as its use in the plural would generate an illicit sequence, i.e., \**goss*.

This kind of interaction between morphological and phonological constraints in allomorph selection is what Paster (2006) labels the "P ≫ M approach", which has its roots in McCarthy & Prince (1993).[7] It is the most widely used model in the analysis of phonologically conditioned allomorphy, and it states that phonological constraints can outrank morphological ones and thus have an impact on the final shape of an affix ('P' stands for phonology and 'M' for morphology). As mentioned, Bonet et al. (2007)'s analysis of Catalan gender exponence is also framed within the P ≫ M approach: it assumes that syllable well-formedness constraints dominate PRIORITY, which is morphological in nature. The P ≫ M model may be necessary to explain other morphophonological phenomena, but one of the purposes of this thesis is to show that gender exponence in Pallarese Catalan can be analyzed without reference to morphological constraints. In this way, phonology and morphology can maintain their independence, thus satisfying the *desideratum* of modularity.

Finally, there is one last thing that needs to be clarified. Until now we have explored what phonology and morphology can or cannot do, but nothing has been said of lexical storage. Any theoretical approach has to deal with cases that deviate from the general pattern and which are not supplied by the generative device. Therefore, the kind of inflectional process defended in this dissertation is valid exlusively for the regular elements of the nominal system. That is, any element which does not bear the default inflectional ending of both masculine and feminine gender is considered to be lexicalized. But what does it mean, exactly, that a structure is 'lexicalized'? The view taken in this work is that the inflectional position (labelled here theme position), which is generally dissociated from the root and has an independent status, is in fact attached to the root in the lexical entries of these non-canonical elements, and the place features associated with it are already spelled out (as opposed to regular nominal elements). Thus, the Vocabulary Items for this set of nouns/adjectives/pronouns are fully specified.

In Pallarese Catalan, -[a] is the default feminine marker, as in *cas*[a] 'house'. However, some feminine nouns bear instead different endings, like -[e] in the case of *mar*[e], shown in (11):

(11)     cas-[a]    'house'

---

mar-[e] 'mother'

Even though *mare* is not realized with the canonical feminine ending, it belongs to the feminine gender nevertheless, and thus the vowel -[e] must be somehow specified in the lexical entry to prevent it from getting the default feminine exponent, i.e., *\*mar*[a]. Therefore, the lexical entries for a regular —(12-a)— and a non-regular noun —(12-b)— look as follows:

(12)   a.   /káz-V/
       b.   /máɾ-e/

The morphological and phonological structures posited for all classes of Pallarese nouns and adjectives will be the topic of 7.1, but as for now the key point is to distinguish between storage and computation. Whereas in (12-a) the theme position is underspecified and consists only of a V-slot, in (12-b) the inflectional marker is fully spelled out. It will be shown in chapter 4 that gender realization is tightly connected to the theme position, which is generally only spelled out by an underspecified vowel, as in (12-a), but, nevertheless, both domains, gender and the morphological theme position, remain independent. In the case of (12-b) there is no connection between the inflectional marker -*e* and gender whatsoever, as opposed to the ending -*a*, which spells out [+fem] in the step that follows the insertion of -V (i.e., a feature [−cons] connected to an X-slot) and links it precisely to this V-slot, thus yielding [káza]. The inflectional theme position in *mar*[e], instead, is fully realized in the lexical entry, and no subsequent steps insert either an underspecified vowel or the featural exponents associated with [+fem]. Everything is, as already mentioned, lexicalized, which implies that this set of words does not enter the regular phonological spell-out process, as opposed to nouns like *casa*.[8] It must be noted, though, that these lexicalized nouns contain structural information, and it is thus clear that -*e* is a theme vowel.[9]

This idea of lexicalized structures contrasts with the view of diacritics specified in the root. This is in fact the solution adopted in other analyses of Catalan (and

---

[8]I make use of the terms 'lexicon' and 'lexicalized structures' in this dissertation to refer to the storage of forms that are not generated regularly, although there is no proper lexicon in DM or other non-lexicalist frameworks. Lexical insertion is still postsyntactic, though.

[9]This is also the idea developed in Bermúdez-Otero (2013, p. 51) for the stem-storage (with theme vowels) of Spanish nouns: "Crucially, an attribute may consist of a complex representation with internal constituent structure, and relationships of correspondence between constituents of different attributes are expressed by coindexation." ('attribute' refers here to the information processed by every separate module, i.e., phonology, morphosyntax, etc.) The morphosyntactic category 'noun' consists thus of a root plus a theme position, and both the root and Th are coindexed with a specific set of exponents but stored as a complex entity.

Spanish) nominal inflection for non-canonical word endings, as well as in Bonet et al. (2007) (see 3.1.1 for more details). The main advantage of diacritics is that we can maintain the 'independence' of roots. That is, the lexicon stores roots, on the one hand, and inflectional sufixes, on the other hand, and it is the morphological component is where inflected words are created by assembling those roots and suffixes. In order to provide the correct ending for non-default nouns, however, roots need to be marked somehow, as in (13) (*mare* 'mother' is used again as an example):

(13)    /máɾ$_e$/

Crucially, the subscript letter in (13) is not an inflectional marker. It is merely a specification of the root to generate the right pattern. But the argument for root diacritics is, in a way, circular. Roots must carry diacritics because they need to surface with the right inflectional marker, and as Acquaviva (2009, p. 2) points out, "positing an invisible class marker on a root in order to make sure that it ends up in the right nominal or verbal inflectional class simply states the observed correlations (if noun, class X, if verb, class Y), treating them as part of the root itself". Why not assume, then, that this information is in fact fully specified, i.e., with structure, in the lexical entry? The disadvantage of this approach is that whole constructions must be stored in the lexicon. That is, diacritics allow the storage of roots in all cases, but if a noun is fully spelled out it means that the morphosyntactic inflectional position is already attached to the root and, thus, it is an inflected word what is stored, not a root. As already stated, though, root diacritics provide information that is not relevant for the syntax (i.e., for agreement) and their only purpose is to provide the right phonological exponent, which, in the end, makes reference still to the inflectional position.[10] The advantage of assuming a lexicalized structure is that there is no need for special 'objects' in roots. The core of the nominal system is still built by the generative device, but all exceptions are instead listed as inflected forms.

The differences between morphosyntactic and phonological computations, on the one hand, and lexical storage, on the other, are thus clear. With this in mind, we can now proceed to the outline of the thesis in 1.2.

---

[10] "There is also empirical evidence that roots should not carry diacritics, in particular class diacritics like declension and conjugation class. Suppose they do: then, these diacritics must be visible in the syntax, in so far as roots occupy syntactic nodes (...) However, this sort of information never enters into relations of syntactic agreement. Not only that: if roots had class features, we would expect that some heads attaching to them should be sensitive to the choice of features. This would lead, for example, to nominalizations only for roots that surface as 2nd declension verbs, or to adjectives corresponding only to nouns with irregular plural, or to causative verbs built only on the roots of -a-stem nouns. But this does not seem to happen." (Acquaviva, 2009, p. 3)

## 1.2   Outline of the thesis

The empirical support for this dissertation is presented in chapter 2, where a full description of the Pallarese Catalan clitic system (both in isolation and combinations), together with the phonological conditionings on the surface forms of the clitics, is given.

Chapter 3 reviews the previous approaches devoted to Catalan and Spanish nominal inflection. The limitations of other analyses and the ideas taken from them for the proposal presented here are discussed, showing that the morphological structure posited by Oltra-Massuet (1999) for the nominal system of Catalan is the most adequate for the analysis of Pallarese clitics.

Chapter 4 develops the analysis of the exponence of gender in the nominal system of Pallarese. The proposal to be presented assumes that every functional syntactic head projects a morphological theme position, which is the location for inflectional exponents. This theme position is spelled out by an underspecified vowel with no place features, whereas gender is realized by floating features that end up linked to the vocalic slot of the theme position under certain conditions. The interactions between epenthesis and theme vowels are also presented in chapter 4, showing that word-final epenthesis is avoided because it interferes with morphological structure. A brief discussion about the location of gender features in the morphosyntactic nominal structure is also provided in this chapter.

Chapters 5 and 6 constitute the core of this dissertation. In chapter 5 the asymmetries regarding gender, number and case found in 3rd person pronominal clitics are discussed and an OT analysis of the data is proposed. Treating the phonological features associated with gender in autosegmental terms allows their use as a repair strategy in phonotactically illicit sequences. This is visible in the alternation *l/lo* of the 3rd person masculine accusative clitic, where the phonological features of the vowel *-o* (the exponent of [–fem]) are attached to the V-slot of the theme position for syllabification purposes. Something similar happens in the 3rd person plural accusative (*ls/les*), although gender is not part of the featural inventory of this clitic due to Impoverishment (Bonet, 1991) and thus no phonological features are present in the input. Nevertheless, the V-slot of the theme position remains and the default phonological features of the unmarked vowel of Pallarese (i.e., *-e*) are inserted. The same holds true for 3rd person plural datives, which share their phonological shape with accusatives. To conclude, the opaque clusters involving 3rd person dative and accusative clitics are also discussed.

Chapter 6 is devoted to the analysis of the rest of Pallarese clitics (also in combinations), showing that the salience of the theme position is also crucial in their

exponence. Since these clitics do not bear gender features, they can only attach the place features of the unmarked vowel *-e* to the vocalic slot provided by the theme position, as in 3rd person plural accusative and dative clitics. 1st and 2nd person plural clitics, which show more problems for interpretation, are also analyzed in this chapter.

In chapter 7 the proposal is extended to nouns and adjectives. Since only *-a* (fem.) and *-o* (masc.) are (floating) exponents of gender, all other inflectional exponents must be specified in the lexial entries of the nouns and adjectives that bear them. Storing full stems that include the theme vowel in non-default cases has the advantage of dispensing with class markers (Harris, 1985) and gender allomorphy in the input (Bonet et al., 2007).

Chapter 8 contains concluding remarks and a comparison between Central Catalan and Pallarese regarding the exponence of clitics. Central Catalan shows an asymmetric pattern in the use of epenthesis, since proclitic and enclitic positions determine whether the clitic undergoes initial or final epenthesis. This situation challenges the theoretical proposal defended in this thesis, which predicts that epenthesis of place features should be preferred whenever there is a theme position available (i.e., initial epenthesis should be avoided). A full account of the Central Catalan clitic system is left for further research.

# TWO

## THE PALLARESE CATALAN CLITIC SYSTEM

## 2.1 General Description

### 2.1.1 Socio-Geographic Information

Pallarese Catalan is a subdialect of Northwestern Catalan, which is spoken in some areas of the Catalan provinces of Lleida and Tarragona, as well as in the eastern region of Aragon bordering Catalonia known as 'La Franja'. All the dialects within this domain share some phonological and morphological features, but the Catalan variety spoken in the Pallars area shows some particularities which motivate its study here as a separate entity.

Pallarese is spoken in the Catalan counties of Pallars Sobirà and Pallars Jussà. Although both are Pyrenean regions, Pallars Jussà is located in a river basin, a fact that makes it more easily accessible. As a result, the dialect spoken here has had frequent contact with other varieties of Catalan and therefore tends to show mixed features, some of them belonging to 'original' Pallarese and others belonging to Standard Catalan (which is mostly based upon Central Catalan, an eastern variety of Catalan). The variant spoken in Pallars Jussà belongs mainly to the 'transition area' of Northwestern Catalan (i.e., an area which has experienced considerable orientalization but still preserves many of its western characteristics), one of the three geolinguistic divisions of the Northwestern linguistic domain detailed in Viaplana (1999).

Pallars Sobirà, by contrast, is located in the heart of the Pyrenees. It has traditionally been isolated from other counties and even today acces to it is not easy. The number of inhabitants has decreased over the last one hundred and fifty years (its population peak was reached in 1860, with about 20,000 inhabitants in the cen-

sus), and today only 7,060 people live there.[1] Consequently, the dialect has been largely unaffected by other Catalan dialects, and it is for this reason that Pallars Sobirà was the area chosen to conduct the study. Nevertheless, in the last twenty-five years the attrition due to the influence of Standard Catalan (as a consequence of the widespread use of radio and television) has been much higher than in Pallars Jussà (as is generally true for the Northwestern Catalan domain as a whole, see Valls 2013).

### 2.1.2 Methodology

Pallars Sobirà has three large valleys: Vall d'Àneu, Vall Ferrera and Vall de Cardós. Travel between them is not particularly easy and the speakers of each valley have their own dialectal specificities, especially regarding lexical items. The data analyzed in this thesis were collected through eleven recorded interviews with local people of Vall d'Àneu and Vall Ferrera who were at the time (May 2014) between the ages of 67 and 88. Five women and six men participated in the study. Three other participants were not included in the analysis either because they could not complete the task or because they were not consistent with the selectional criteria. Speakers were native to the following villages of Pallars Sobirà (their parents having been born there too):[2]

- Vall Ferrera: Ainet de Besan (1), Farrera (1), Araós (2), Àreu (3).

- Vall d'Àneu: Alós d'Isil (1), Estaís (1), Son (1), València d'Àneu (1).

The population of these villages ranged roughly between 10 and 120 inhabitants. These are estimations given by the inhabitants, as the census does not reflect the real demographics of the area (many non-locals are registered in the villages because they have properties there but are actually living elsewhere). Younger speakers were excluded from the sample because of their considerable exposure to mass media as well as greater mobility.[3] Most of the speakers interviewed had never received any education in Catalan and had dropped out of school (taught in Spanish) at early ages, even before finishing high school.[4] Despite there being a gap of twenty years

---

[1]Population statistics as by 2015 (http://www.idescat.cat/emex/?id=26).

[2]A full description of each speaker can be found in appendix A.

[3]Younger generations of Catalans have been educated –totally or partially– in Standard Catalan, whether at school or at university. Moreover, many of them have spent long periods of time away, particularly in the larger provincial towns or in Barcelona. This makes them more vulnerable to changes in their speech.

[4]During the short period of the Spanish Second Republic (1931-1936), just before the Spanish Civil War (1936-1939), Catalan was allowed to be taught at school. Some of the older speakers

separating older and younger speakers, no significant age differences were found in their dialects. However, big differences were found between male and female speakers, men being much more conservative than women (in fact, the youngest informant, a 67-year-old man, had preserved more dialectal features than most of the older female speakers, aged 85 to 88). This is not surprising, as according to Labov (1990) innovations are female-directed. A possible explanation for this in the context of Pallars Sobirà is women's greater exposure to mass media. Although both women and men participate in farming tasks, women are also housewives, giving them more time in the home and therefore exposure to television or radio. By contrast, men spend most of their time outside the household, quite often socializing with other men in environments where the mass media are much less likely to be present.[5] The mass media in Catalonia are almost exclusively dominated by Standard Catalan (it is difficult to find dialectal variants in them), which is considered the prestige register of the language. Therefore, "men use more nonstandard forms, less influenced by the social stigma directed against them; or, conversely, women use more standard forms, responding to the overt prestige associated with them" (Labov 1990). These differences fit into what Labov (1990) has called Principle I and Principle II in the linguistic differentiation between men and women:

(I) In stable sociolinguistic stratification, men use a higher frequency of nonstandard forms than women.

(II) In the majority of linguistic changes, women use a higher frequency of the incoming forms than men.

Although from a sociolinguistic perspective these findings are very interesting, this is not the focus of this dissertation. Gender differences, as well as other sociological factors, will not be part of the analysis.

---

might therefore have had some schooling in Catalan, but their exposure to Standard Catalan must have been minimal. Under the Franco dictatorship (1939-1975) Catalan was strictly forbidden at school, as well as in all other areas of pblic life.

[5]There is a vigorous debate in sociolinguistics literature about the role the mass media play in linguistic change (see Sayers 2014 for a proposal on how to incorporate media factors in sociolinguistic research). Although Trudgill (2014) thinks that the media do not have any influence in changing core lingustic features, I think it is difficult to find an alternative explanation for Pallars Sobirà speakers' behavior. Changes in the use of clitic forms as a consequence of face-to-face interactions with Central Catalan speakers seems to be implausible, as the participants surveyed had never had prolonged exposure to other dialects. The role of the media in shaping urban English speakers' language may be minimal because of a whole suite of internal and external factors, but the Catalan-speaking community of Pallars Sobirà differs substantially from these speakers. Pallarese speakers are exposed to only one variety in the media (as opposed to different varieties in English) and, moreover, their contact with Standard Catalan comes mainly through television and radio. For these reasons, I think the influence of mass media is the most likely culprit for the changes observed.

The interviews consisted of a questionnaire about clitic configurations (either in isolation or in two-clitic combinations) in which participants had to translate sentences from Spanish into Catalan.[6] A priori, this method does not seem to be the most appropiate, as Spanish could interfere with Catalan in the elicitation. Moreover, speakers might tend to use a higher register of Catalan (i.e., Standard Catalan), which would imply a sharp deviation from their colloquial speech (as Standard Catalan is not based on Northwestern varieties). However, this method was used because it is impossible to elicit all clitic configurations in a casual speech context, as Catalan has a particularly complex clitic system. In some speakers the translation task caused interferences, but in general it seemed to have little impact on the final output of the speaker.

In order to create an informal environment appropriate for casual speech, participants were asked to talk about their lives before and/or after each interview. As Labov (1972) pointed out, casual speech outside of the formal interview are crucial to collect useful linguistic data. The topics raised in the conversations were the following:

- childhood memories (particularly, memories of the Spanish Civil War).

- daily duties (farming and agricultural tasks).

- changes over the years in the natural and cultural landscape of the speaker's valley.

Interviews lasted between 35 and 90 minutes, being on average around 50 minutes in duration, with a short break in the middle. During this pause, pictures of animals, objects and plants were shown to the participants and they were asked to name them. The lexical units referring to these items were specific to the area, so the exposure to them encouraged the use of genuine dialectal expressions.

Few studies have analyzed the phonology and morphology of Pallarese. To my knowledge, only Coromines (1936), Sistac (1998), Viaplana et al. (2007), Alcover & Moll (2011) and Massanell (2012) have studied Pallarese in depth. The problem with most of these materials is their lack of information regarding the particular villages surveyed in this study, as the recordings and data come mainly from bigger villages like Rialp, Llavorsí or Sort (communities which are actually not in the

---

[6]The questionnaire can be found in appendix B. Ideally, every speaker should have elicited all clitics in all contexts. In practice, this would have taken more time than seemed appropriate given the advanced age of the participants. For this reason, whenever a group of clitics had a shared behavior, only one of them was asked for. Confrontation with casual speech filled the gaps that were left empty after the translation task.

valleys). Nonetheless, some of the data from these dialectological sources coincide with what can be seen in the interviews conducted in Vall d'Àneu and Vall Ferrera. However, to the best of my knowledge the remainder of the data collected here has not been examined in previous research. Thus, the data presented in the rest of this chapter aim to fill a gap in the morphophonological study of Northwestern varieties of Catalan (and Catalan as a whole).[7]

## 2.1.3 Morphophonological description

Like other Catalan dialects and Romance languages, Pallarese Catalan has a rich clitic system made up of fourteen morpho-syntactic units. Table 2.1 shows the full clitic system in isolation:

| Label | Citation form | Contextual variants |
|---|---|---|
| 1st sg. Acc./Dat. | me | m', 'm, me |
| 2nd sg. Acc./Dat. | te | t', 't, te |
| 1st pl. Acc./Dat. | mos | mos, mo |
| 2nd pl. Acc./Dat. | vos | vos, us, vo |
| 3rd Acc. sg. masc. | lo | lo, l' |
| 3rd Acc. sg. fem. | la | la, l' |
| 3rd Acc. pl. masc | les, es | les, 'ls, 's |
| 3rd Acc. pl. fem. | les, es | les, 'ls, 's |
| 3rd sg./pl. reflex. | se | s', 's, se |
| 3rd Dat. sg. | li | li |
| 3rd Dat. pl. | les (+hi) | les (+hi) |
| Neuter | ho | ho |
| Partitive | ne | n', ne |
| Locative | hi | hi |

Table 2.1: *Pallarese pronominal clitics*

Pronominal clitics can appear as enclitics after an infinitive, gerund or imperative. An example of an imperative + enclitic sequence is shown in (1):

(1)    Agafa-   **la**!
       take.IMPR ACC.3S.FEM

---

[7]Unless specified in the examples, all the data shown in this chapter come from the interviews carried out in Pallars area.

'Take her!'

Proclitics are found after other forms of the verb.[8] A complete description of the clitic system in a 'regular' proclitic preconsonantal position is offered below (a thorough description of all the phonologically conditioned forms will be presented in section 2.2).[9] [10]

- *1st person singular*

  1st and 2nd person clitics express person and number. They have different syntactic functions, most often in argument position. There is case syncretism between dative and accusative —see 2.1.4.

  The regular form in Central Catalan is *em*, *me* being common mainly in Northwestern Catalan (as well as in other varieties).

  (2)  a.  **Me**  sembla  correcte
           DAT.1S seem.3S right
           'It seems right to me'

       b.  **Me**  busquen
           ACC.1S search.3PL
           'They are looking for me'

- *2nd person singular*

  Following the pattern of the 1st person singular, *te* is preferred instead of the regular form *et* of Central Catalan.

  (3)  a.  No **te**   pagaven
           NEG DAT.2S pay.3PL.IMPF
           'They didn't pay you'

       b.  **Te**  seguix  la      policia
           ACC.2S follow.3S the.FEM.S police.FEM.S
           'The police are following you'

- *1st person plural*

  This also differs from Central Catalan, where *ens* is the regular form.

---

[8]Except for participles, which can never host a clitic.

[9]I will use an adapted Catalan orthography to reflect the dialectal differences found in Pallarese. All the examples —sometimes simplified or modified for the sake of clarity— are taken from the interviews recorded in Pallars Sobirà.

[10]Some of the examples reflect casual speech, but most of them are excerpts from the Spanish-to-Catalan translation task. In some instances, the examples provided in this chapter do not provide sufficient context for the syntactic functions of the clitics to be understood. However, they are not ambiguous in the discourse context in which they occurred.

(4) a. **Mos** criden per anar al   cine
     ACC.1PL call.3PL for go.INF to+the.MASC.S cinema.MASC.S
     'They are calling us to go to the cinema'

   b. Lo   lluç  **mos**  agrada bullit
     the.masc.s hake.MASC.S DAT.1PL please.3S boil.PART
     'We like to eat our hake boiled'

- *2nd person plural*

  The standard form in Northwestern Catalan is *vos*, whereas Central Catalan uses the form *us*. *Mos* and *vos* share the inflectional ending *-s*, a marker of plurality in Catalan. Alternatively, one could argue that the clitic as a whole is expressing plurality, in which case *-s* would not be a separate morph (see section 6.2 for more details).

(5) a. **Vos**  necessita
     ACC.2PL need.3S
     'S/he needs you'

   b. **Vos**  dóna la   mà
     DAT.2PL give.3S the.FEM.S hand.FEM.S
     'S/he shakes your hand'

- *3rd person feminine accusative singular*[11]

  3rd person singular accusative clitics reflect gender overtly. The default exponent for feminine in Catalan is *-a*, and this is what we find in the clitic:

(6) La   paraula **la**   pronuncies d'una
   the.FEM.S word.FEM.S ACC.3S.FEM pronounce.2S of+a.FEM.S
   forma
   way.FEM.S
   'You pronounce the word in a certain way'

- *3rd person masculine accusative singular*

  The marked exponent for masculine in Catalan is *-o* (*-Ø* being its unmarked counterpart). It appears with some nouns and adjectives, and also in the clitic:

(7) Lo   llit  **lo**   dixaré  net
   the.MASC.S word.MASC.S ACC.3S.MASC leave.2S.FUT clean.MASC.S
   'I will leave the bed clean'

---

[11]The phonological shape of 3rd person accusative clitics is shared with the definite article, in both singular and plural. This syncretism is analyzed in section 2.3.

- *3rd person feminine accusative plural + 3rd person masculine accusative plural*

  3rd person plural accusative clitics are syncretic, and both masculine and feminine end up taking the same shape (see 2.1.4). There are two allomorphs, *les* and *es*, which are in free variation. (8-a) and (9-a) show how the clitic *les* can substitute for feminine or masculine antecedents, and the same holds for *es* in (8-b) and (9-b):

  (8)   a.   Les          tendes       **les**          embruten        ells
             the.FEM.PL tent.FEM.PL ACC.3PL.FEM make-dirty.3PL they
             'They make the tents dirty'

        b.   Es           tomates      **es**           fiques a l'
             the.FEM.PL tomato.FEM.PL ACC.3PL.FEM put.2S to the.MASC.S
             enciam
             lettuce.MASC.S
             'You put the tomatoes in the lettuce'

  (9)   a.   Les          llits        **les**          embruten        ells
             the.MASC.PL bed.MASC.PL ACC.3PL.MASC make-dirty.3PL they
             'They make the beds dirty'

        b.   Les          més petits   **es**           baixen
             the.MASC.PL more small.MASC.PL ACC.3PL.MASC go-down.3PL
             amb cotxe       escolar
             with car.MASC.S school.MASC.S
             'The younger children travel by schoolbus'

- *3rd person reflexive*

  The label 'reflexive' can sometimes be confusing because *se* reflects different uses, like the impersonal (as in (10-a)), but I will adopt it here for the sake of convenience.[12] It is used for the 3rd person singular and plural.

  (10)  a.   La           llet         no **se** cotizava          bé
             the.FEM.S milk.FEM.S NEG IMP contribute.3S.IMPF well
             'We weren't getting paid much for milk'
        b.   Es           meus         amics        **se**
             the.MASC.PL my.1PL.MASC friends.MASC.PL REFL.3PL
             volen      casar
             want.3PL marry.INF
             'My friends want to get married'

---

[12]See Todolí (2002) for a complete review of all the syntactic functions that clitics can express.

    c.    Aguesta    gent    **se**    cansa    molt de
        this.FEM.S people.FEM.S REFL.3S get-tired.3S very of
        pressa
        hurry.FEM.S
        'These people get tired very quickly'

- *3rd person dative singular*

3rd person singular clitics differ in their phonological shape for dative and accusative. The form *li* expresses dative singular in Pallarese:

(11)    **Li**    compres un    regalo
        DAT.3S buy.3S    a.MASC.S present.MASC.S
        'You buy a present for her/him'

- *3rd person dative plural*

The dative plural is expressed by the clitic *les*, which is exactly the same shape taken by the accusative plural clitics. In some cases, it is reinforced with *-i* (orthographically *hi*), as in the dative singular.[13]

(12)    a.    **Les**    cantes una    cançó
            DAT.1PL sing.2S a.FEM.S song.FEM.S
            'You sing them a song'

        b.    **Les**    **hi**    agafo la    mà
            DAT.1PL DAT take.1S the.FEM.S hand.FEM.S
            'I take their hand'

- *Neuter*

The clitic *ho* usually replaces neuter pronouns, noun phrases or embedded clauses. It can also substitute for a predicative expression or be the inherent clitic of some verbs, although these are much less frequent scenarios.

(13)    No **ho**    sé
        NEG NEUT    know.1S
        'I don't know that'

- *Partitive*

In most instances *ne* expresses partitivity, either as a subject or as a complement of a verb, although it is quite common to find it having locative uses. It

---

[13]*Es* is also a possible allomorph —as in the 3rd person plural accusative; see section 2.2.

can also work as a predicative expression or inherent clitic. It differs from the regular form *en* in Standard Catalan:

(14)    **Ne** vull    quatre
      PART want.1S FOUR
      'I want four of those'

- *Locative*

  The pronoun *hi* usually works as the locative complement of the verb, but sometimes it performs other syntactic functions such as adjunct of the predicate or inherent clitic:

(15)    **Hi** anàvom
      LOC go.1PL.IMPF
      'We used to go there'

### 2.1.3.1   Combinations

Clitics can also combine with one another, and combinations of two clitics are in fact very frequent (I leave aside combinations of more than two clitics, which will not be discussed). An example of a clitic sequence is presented below:

(16)    **T'**    **ho**    compro
      DAT.2S NEUTER buy.1S
      'I'll buy this for you'

Clitic sequences follow a fixed pattern, that is, not all combinations are possible and the ordering is not free. Several restrictions (which apply to the whole domain of Catalan) need to be taken into account:[14]

1. Same clitic combinations are not allowed. Although syntactically two different functions can be expressed with the same clitic, an identical sequence is forbidden.

2. Catalan is subject to the Person Case Constraint —or PCC— (Bonet 1991), which states that combinations of 1st and 2nd person accusative + 3rd person dative clitics are not allowed (examples taken from Bonet 2008b):

   (17)    *Al    director    **me**    **li**    ha recomanat
          to+the director.MASC.S ACC.1S DAT.3S has recommended
          la      Mireia
          the.FEM.S Mireia

---

[14]See Bonet (2002) for a complete review of the restrictions on clitic combinations.

'As for the director, Mireia has recommended me to him'

If the order is reversed, that is, if the first clitic is a 1st or 2nd person dative and the second clitic is a 3rd person accusative, the sequence is totally acceptable:

(18)    El          director,        **me**    **l'**    ha  recomanat
        the.MASC.S director.MASC.S DAT.1S ACC.3S has recommended
        la          Mireia
        the.FEM.S Mireia
        'Mireia has recommended me to the director'

This also happens with 1st and 2nd person pronouns (second person precedes first person). Although rare and acceptable only for some speakers, they can appear in sequences like the following:

(19)    **Te**    **m'**    ha  recomanat    la          Mireia
        ACC.2S DAT.1S has recommended the.FEM.S Mireia
        'Mireia has recommended me to the director'

As a consequence of this, two different versions of the PCC have been proposed:

- Strong version: the direct object must be a third person.

- Weak version: if there is a third person, it must be a direct object.

3. 3rd person clitic combinations are not phonologically transparent. In Pallarese, these are the outputs for the combination of 3rd person clitics:

(20)

| li (dat.sg) | + | la (acc.fem.sg) | = | l'hi | [li] |
|---|---|---|---|---|---|
| li (dat.sg) | + | lo (acc.masc.sg) | = | l'hi | [li] |
| li (dat.sg) | + | les (acc.fem.pl) | = | les hi | [lezi] |
| li (dat.sg) | + | les (acc.masc.pl) | = | les hi | [lezi] |
| les (dat.pl) | + | la (acc.fem.sg) | = | les hi | [lezi] |
| les (dat.pl) | + | lo (acc.masc.sg) | = | les hi | [lezi] |
| les (dat.pl) | + | les (acc.fem.pl) | = | les hi | [lezi] |
| les (dat.pl) | + | les (acc.masc.pl) | = | les hi | [lezi] |

In single clitics, the -*i* marker of the dative plural is not very common, but in combination its appearance is much more common (it is the default, although the cluster sometimes sufaces without -*i*, i.e., with the shape of the dative plural in isolation; see 5.1.3 for more details).

4. The clitic *ho* presents some peculiarities too. In combination with *li*, it also shows up as *l'hi*:[15]

   (21)  Açò          **l'hi** dóna  a sa      germana
         this.NEUTER ho+li give.3S to POSS.3S sister.FEM.S
         'S/he gives this to her/his sister'

   For some speakers (crucially, those preserving the greatest number of dialectal features) the combination yields a different sequence, *la hi*:

   (22)  Això,         ensenya- **la-hi** d'una       vegada
         this.NEUTER show.2IMP ho+li of+a.FEM.S time.FEM.S
         'Show this to her/him now.'

   This is very atypical for Catalan, as it seems that a feminine clitic *la* is used to express the neuter. To my knowledge, this combination is not possible either in Western varieties or in Eastern dialects. However, this was reported in the data collected by Antoni M. Alcover during his fieldwork in the Catalan-speaking territories between 1900 and 1928 (a material that has been digitalized in Alcover & Moll 2011).

   (23)  dóna**lahi** = dona **axò** a ell 'Give it to him = Give that to him'[16]

   Although we have no audio record of this form to confirm it, *lahi* would seem to be pronounced [laj]. That is, a neuter pronoun being substituted by a feminine clitic, which is a sharp deviation indeed from the regular uses of Catalan.

   When *ho* appears in combinations with the dative plural *les (hi)*, the output is also *les hi*, as in 3rd person accusative + dative combinations in (20) (e.g., *Això, als teus fills,* les hi *dono* 'I give this to your kids').

5. In many Catalan dialects, the combination *li + en* yields the sequence *n'hi*. In Pallarese Catalan, however, it is acceptable:

   (24)  **Li'**   **n**  donaré      quatre, de roses
         DAT.3S part give.1S.FUT four     of roses.FEM.PL
         'I will give him four roses'

---

[15]Mascaró (1986) considers that the pronoun *ho* has been substituted by the 3rd person accusative masculine singular pronoun /l/ in these combinations, but see section 5.1.3 for other interpretations.

[16]This is an exact reproduction of the transcription given by Antoni M. Alcover in his fieldwork notebooks. In this one instance, the orthography differs from what is generally used here.

6. Other combinations, like *en + ho* and *ho + hi*, are rare in the whole domain of Catalan and are usually avoided by using a different combination of pronouns or by not producing them at all. As for Pallarese Catalan, they could not be elicited during the interviews (see section 6.3 for more details).

As shown in the examples above, clitics follow a fixed order when they appear in combinations. This is the general scheme proposed by Bonet (1995b) for Standard Catalan:

(25)    *es* - 2nd - 1st - 3rd.dat.pl - 3rd.acc - genitive - oblique/neuter 3rd.dat.sg

The fact that the 3rd dative plural and the 3rd dative singular appear in different positions is due to the changes in clitic combinations found in dative + accusative sequences. In Standard Catalan we find 3rd person dative plural + 3rd person accusative (**Els la** *compro*), but also (in fact, this is the default option in Central Catalan) the combination *l'hi* (**L'hi** *compro*), which seems to correspond to a sequence of 3rd person accusative singular + 3rd person dative singular —although this combination is obscure and it is thus unclear which segments correspond to each clitic. In Pallarese Catalan, only the order 3rd person accusative + 3rd person dative is found, so the scheme should be modified as follows:

(26)    *se* - 2nd - 1st - 3rd.acc - 3rd.dat - gen. (*ne*) - oblique (*hi*)/neuter (*ho*)

Following this pattern, we can generate all the structures previously exemplified: *se t'escapa (*te s'escapa), les hi compro (*li'ls compro), n'hi ha (*hi n'ha), t'ho dóna (*ho't dóna)*, etc.

These are the basic facts regarding clitic combinations in Pallarese Catalan. The combinations that involve 3rd person clitics are analyzed in chapter 5, while all other clitic combinations are analyzed in chapter 6.

### 2.1.4 Syncretism in Pallarese Clitics

Romance languages tend to show syncretism in accusative and dative clitics (see Nevins & Savescu 2010 for an analysis of this kind of syncretism in Romanian). As previously shown for Pallarese, 1st and 2nd persons share their form in accusative and dative (in both singular and plural), something true also for Spanish and Italian, as illustrated in (27) and (28) below:

(27)    a.    **Te** compro un     libro
             DAT.2S    buy.1S a.MASC.S book.MASC.S
             'I'll buy a book for you'

b. **Te** llevo a casa
ACC.2S bring.1S to home.FEM.S
'I'll take you home'

(28) a. **Ci** offre un caffè
DAT.1PL offer.3S a.MASC.S coffe.MASC.S
'S/he buys us a coffe'

b. **Ci** porta al mare
ACC.1PL bring.3S to-the.MASC.S sea.MASC.S
'S/he takes us to the sea'

We see a different, less common kind of syncretism in 3rd person accusative plural clitics. The examples given in section 2.1, repeated in (29), show that the masculine and feminine end up taking the same phonological form:

(29) a. Les tendes **les** embruten ells
the.FEM.PL tent.FEM.PL ACC.3PL.fem make-dirty.3PL they
'They make the tents dirty'

b. Les llits **les** embruten ells
the.MASC.PL bed.MASC.PL ACC.3PL.MASC make-dirty.3PL they
'They make the beds dirty'

Although in the singular there are two different clitics, one for the masculine, *lo*, and one for the feminine, *la*, there seems to be a neutralization in the plural, with *les* (or *es*, its allomorph) being its only phonological outcome. This phenomenon is peculiar to Pallarese, since other Catalan dialects like Central Catalan maintain the gender distinction in the plural clitic, as shown in (30):

(30) a. Les tendes **les** embruten ells
the.FEM.PL tent.FEM.PL ACC.3PL.FEM make-dirty.3PL they
'They make the tents dirty'

b. Els llits **els** embruten ells
the.MASC.PL bed.MASC.PL ACC.3PL.MASC make-dirty.3PL they
'They make the beds dirty'

Some languages, like French, display the same behavior as Pallarese (i.e., they have a gender distinction only in the singular clitic):

(31) a. Je **le** vois
1S ACC.3S.MASC see.1S
'I see him'

b. Je **la** vois
1S ACC.3S.FEM see.1S
'I see her'

> c.   Je **les**               vois
>      1s ACC.3PL.MASC/FEM see.1s
>      'I see them' (men or women)

Moreover, this syncretism is only active in the clitic (as well as in the definite article, see section 2.3), because the differences between masculine and feminine are still visible in the rest of the nominal system:

(32)   a.   Llits         petits
            bed.MASC.PL small.MASC.PL
            'Small beds'

       b.   Tend**es**       petit**es**
            tent.FEM.PL small.FEM.PL
            'Small tents'

In (32-a) and (32-b) we find -*Ø* and -*a* (-*e* due to a phonological process) as gender markers, the two markers being the default masculine and feminine exponents, respectively. But as shown by the data, this is not what happens in the case of the clitic, which always remains *les*. It is clear, thus, that there is an asymmetry in the clitics' gender exponence regarding number, which will be analyzed later in section 5.1.2.

## 2.2   Phonological Conditionings on Clitics' Surface Forms

The surface form of clitics is phonologically conditioned. Prosodic requirements (together with some morphological restrictions) determine the final outcome of the clitic, and this is the reason why we find contextual variants of the same pronoun.

As shown in the previous section, pronominal clitics can appear either isolated or in combinations, and their phonological shape varies accordingly. In this section I illustrate all the phonological conditionings that affect clitics. Some processes apply only to certain kinds of clitics, so they have been grouped according to their behavior.

### 2.2.1   Clitics in Isolation

- *1st/2nd person singular, neuter and partitive*

  *Me*, *te*, *se* and *ne* display the same characteristics when exposed to certain phonological environments. Two contextual variants are found in Pallarese

Catalan, depending on whether or not they can be attached to a preceding/following vowel:

| _C | C_C | _V | V_V | C_V | V_C |
|------|------|-----|-----|-----|-----|
| [me] | [me] | [m] | [m] | [m] | [m] |
| [te] | [te] | [t] | [t] | [t] | [t] |
| [se] | [se] | [s] | [s] | [s] | [s] |
| [ne] | [ne] | [n] | [n] | [n] | [n] |

In preconsonantal position, the full form is chosen:

(33)  **[te]**  vull  dir  açò
      DAT.2S want.1S say.INF this.NEUT
      'I want to tell you this'

A complex onset like *[tb] is not permitted in Catalan due to the Sonority Sequencing Principle (Clements, 1990), so the full form is chosen. When preceded by a consonant, the same strategy is used:

(34)  Aguestes  iclésies  **[me]**  semblen  molt
      this.FEM/PL church.FEM.PL DAT.1S seem.3.PL very
      boniques
      beautiful.FEM.PL
      'These churches seem very beautiful to me'

The cluster *[sm] is forbidden as a complex coda in Catalan due to its rising sonority, and the full form is thus preferred (see Lloret 2002 for a complete review of the phonotactics of Catalan).

When there is a vowel next to the clitic, whether before or after it, syllabification is possible and the non-vocalic form is chosen:

(35)  a.  Per això'  **[s]**  va  plegar
          for  this.NEUT IMP.3S go.3S stop.INF
          'This is why we stopped working'
     b.  **[n]**'  agafa  quatre
          PART take.3S four
          'S/he takes four of that'

This is not true in all cases, however. Some clitics are not syllabified with the preceding vowel, thus yielding sequences like *no **se** cotizava* 'it didn't bring

any money' (we would expect *no's cotizava*).[17] Speakers use both indistinguishably, although it seems that *ne* and *se* are more easily contracted than *me* and *te*.

A singular characeristic of Pallarese Catalan, together with all Northwestern Catalan varieties, is the maintenance of *r*-deletion in infinitives even when a clitic is attached. This is a significant difference with regard to Central Catalan, as shown in (36):[18]

(36)

| | | | | | | |
|---|---|---|---|---|---|---|
| **C. Cat.** | *cantar* | [kəntá] | 'to sing' | *cantar-te* | [kəntártə] | 'to sing to you' |
| **Pallarese** | *cantar* | [kantá] | 'to sing' | *canta't* | [kantát] | 'to sing to you' |

In Central Catalan, *r*-deletion is blocked by the adjunction of the clitic. But Pallarese uses a different strategy. The morph *-r* does not surface in any infinitival context and, since the clitic can be syllabified with the theme vowel (i.e., a regular prevocalic context), the contracted form is attached to the verb.[19]

- *3rd person accusative singular*

  Table 2.1 shows that *lo* (masculine) and *la* (feminine) are the regular forms for 3rd person accusative singular clitics.

  If followed by a consonant, nothing changes, as seen in examples (6) and (7). However, if there is a vowel after the clitic, both the masculine and feminine gender markers, *-o* and *-a*, are not present in order to avoid a hiatus.

(37)  a.  La        tenda      [**l**]'        embruten      ells
          the.FEM.S tent.FEM.S ACC.3S.FEM make.dirty.3PL they
          'They make the tent dirty'

      b.  Lo        llit       [**l**]'        embruta       ell
          the.MASC.S bed.MASC.S ACC.3S.MASC make.dirty.3S he
          'S/he makes the bed dirty'

---

[17]Coromines (1936, p. 57) also notes the variability in the appearance of *-e* in this set of clitics in postvocalic contexts.

[18]*r*-deletion is ubiquitous in Central Catalan infinitives —in isolated forms and before a vowel or a consonant— except in the presence of a clitic.

[19]It is unclear whether infinitival *-r* should be present underlyingly at all, as it never surfaces (I leave aside second conjugation verbs like *caure*, *beure*, etc.).

However, when preceded by a vowel, the clitics differ in their final shape. In the feminine, the gender marker is maintained, whereas in the masculine the exponent *-o* is absent:

(38)    a.    No  **[la]**    vull    fer,    aguella    cosa
                NEG ACC.3S.FEM want.1S do.INF that.FEM.S thing.FEM.SG
                'I don't want to do that'

        b.    Lo    llit    que' **[l]**    gafe ell
                the.MASC.S bed.MASC.S take 3S.SUBJ he
                'S/he should take the bed'

According to the data, there seems to be an alternation in the gender exponence of the masculine clitic. Before a consonant, the full form is preferred, as complex onsets can never start with a liquid in Catalan (*[l p]are - [lo] pare, *the father*). But whenever it can be syllabified with another vowel, the non-vocalic form is chosen. The most widespread ending for the masculine is, by far, -Ø, and preference is given to it whenever possible. If a phonotactic problem arises, a gender marker that is less frequent —but also present in the paradigm— is used to solve the phonological conflict. It is important to notice here the differences between Central and Northwestern Catalan in their use of repair strategies. The complex onset mentioned above is illicit for both dialects, but Central Catalan does not use an inflectional marker to avoid it. Instead, regular epenthesis takes place at the left edge of the definite article ([ə]l pare). In other words, the solution is phonological. By contrast, the solution in Northwestern Catalan is morphological (i.e, the use of an inflectional exponent).[20]

These are, then, the forms found in the 3rd person accusative singular:

|  | Before a vowel | Before a consonant | After a vowel |
|---|---|---|---|
| **Masculine** | [l] | [lo] | [l] |
| **Feminine** | [l] | [la] | [la] |

- *3rd person accusative plural*

As discussed in 2.1.4, one of the most singular characteristics of Pallarese is the syncretism found in 3rd person accusative plural clitics. The gender distinction is lost and a single form is used for both masculine and singular,

---

[20]See 8.2 for a comparison of Pallarese and Central Catalan with regard to this kind of epenthetic phenomena.

*les*. For this reason, in a sentence like *No **les** vull* 'I don't want them', *les* can make reference to a masculine or a feminine plural antecedent.

The allomorph *es* seems to be in complete free variation. Speakers use both indistinguishably, even within the same sentence, and one or the other can appear in any context:

(39) Es        tomates        [**es**]/[**les**]  fiques a l'
     the.FEM.PL tomato.FEM.PL  ACC.3PL.FEM put.2S to the.MASC.S
     amanit
     salad
     'You put the tomatoes in the salad'

When the clitic is followed by a vowel, the full form (full = *les/es*) is found:

(40) Les        llits          [**lez**]           embruten       ells
     the.MASC.PL bed.MAASC.PL  ACC.3PL.MASC make.dirty.3PL they
     'They make the beds dirty'

But in cases where there is a preceding vowel, only *-ls* surfaces, followed by either a vowel or a consonant:[21]

(41)  a.  Que' [**lz**]              agafen        ells
          that  ACC.PL.MASC/FEM take.3PL.SUBJ they
          'That they take them'

      b.  Que' [**ls**]              compren       ells
          that  ACC.PL.MASC/FEM buy.3PL.SUBJ they
          'That they buy them'

Yet even in such cases, the variation between *es/les* is also free. However, if *es* is the allomorph chosen, there is vowel deletion to avoid a hiatus (a widespread phenomenon in the phonology of Catalan):

(42) Es        tomates        no' [**s**]       comprem ací
     the.FEM.PL tomato.FEM.PL NEG ACC.3PL.FEM buy.1PL  here
     'We don't buy the tomatoes here'

The sequence *no + es* is phonetically produced as [nos]. This is normal for Catalan, but it is somewhat surprising that the accusative plural clitic ends up just being *-s* (*les > es > s*), partly because it is also the regular plural marker

---

[21]The alternation between [z] and [s] in both *les* and *es* can be attributed to any of these three regular phonological processes of Catalan: voicing assimiliation, sonorization of fricatives or final devoicing.

and partly because there exists another allomorph that would avoid a hiatus while being faithful at the same time (i.e., *les*). That is, [l] expresses 3rd person and [s] expresses plurality, but in this case the result seems to be the fusion of the two morphemes. Moreover, if gender were expressed somehow in the plural clitic (in which case one could not assume that gender was 'lost' during the morphological derivation), three would be then the morphemes fused into [s]: masc/fem + 3rd person + plural.

- *3rd person dative plural*

  The 3rd person dative plural has the same form and almost the same behavior as the 3rd person accusative plural. The basic form is *les*, but an allomorph *es* is also possible, the only difference being that *es* is less frequent than in the accusative. These are some examples of its use:

  (43)  a.  **[les]**  toques  la  mà
        DAT.3PL touch.2S the.FEM.S hand.FEM.S
        'You touch their hand'

        b.  Cosís-  **[ez]**  lo  botó!
        sew.IMPR DAT.3PL the.MASC.S button.MASC.S
        'Sew the button for them!'

  Likewise, the dative takes the contracted form *-ls* before a vowel:

  (44)  Canta'  **[lz]**  una  cançó!
        sing.IMPR DAT.3PL a.FEM.S song.FEM.S
        'Sing them a song!'

- *3rd person dative singular*

  The dative singular always surfaces as *li* in isolation, regardless of the phonological context (before or after a vowel, before or after a consonant):

  (45)  a.  **[li]**  agafa  la  mà
        DAT.3S take.3S the.FEM.S hand.FEM.S
        'S/he takes her hand'

        b.  A la  Maria **[li]**  dem  tres roses
        to the.FEM.S Mary  DAT.3S give.1PL three rose.FEM.PL
        'We give Mary three roses'

- *1st and 2nd person plural*

  1st and 2nd person plural clitics are quite unstable in their phonological shape.

They vary depending on the context and, although they can be treated as a group, there are some differences among them.

The full forms *mos* and *vos* are found before vowels or consonants, as shown above in examples (4) and (5). The main alternation is between [z] and [s], which depends on the voicing quality of the following segment, but as previously mentioned, this is one of the regular phonological processes of Catalan.

There are no differences in the 1st person pronoun if it is preceded by a vowel. That is, other clitics can be attached to a vowel and, thus, alter their phonological shape, but this is not the case for *mos* (cf. (47) and (48)):

(46)  a.  Les           que [**moz**] van     acabar  de xafar
          ACC.3PL.FEM that ACC.1PL go.3PL finish.INF of crush.INF
          'Those things that finally crushed us'

      b.  No [**moz**]   interessa la        política
          NEG DAT.1PL interest.3S the.FEM.S politics.FEM.S
          'We're not interested in politics'

The same does not hold for the 2nd person plural clitic. Two allmorphs are found for it, *vos* ([bos]) and *us* ([wz]).[22] The variant *vos* appears in pre- and postconsonantal environments:

(47)  a.  Estan vigilant-  [**bos**]
          be.3PL guard.GER ACC.2PL
          'They are watching you'


      b.  [**bos**] quedareu     aquí
          INH.2PL stay.FUT.2PL here
          'You'll stay here'

When preceded by a vowel, however, the allomorph *us* is chosen:[23]

(48)  No [**wz**]      allunyeu!
      NEG REFL.2PL get.away.2PL
      'Don't go away!'

Because it is syllabified with the preceding vowel, the contact originates a glide.

---

[22]In certain contexts the approximant [β] surfaces instead of the plosive [b], but this is just due to the widespread process of spirantization found in Catalan.

[23]There is inter- and intraspeaker variation, and sometimes the form *vos* appears in contexts where we should expect *us*, as in *No **vos** nececessita* 'S/he doesn't need you'. Postvocalically, though, *us* seems to be the preferred form, and this is in fact what Coromines (1936, p. 57) reports.

This is the reason why we find the phonetic forms [wz]/[ws] (depending on the following segment) instead of [uz]/[us]. In addition, the inflectional suffix for 2nd person plural in the verbal paradigm is [w] (Mascaró 1986), with a glide like the clitic, and therefore examples like (49) pose problems for segmentation:

(49)    Agafe    [**ws**]!
        hold.2PL REFL.2PL
        'Hold yourselves!'

The verb in isolation is *agafe*[w]. If the clitic is attached to it, either the two morphs are fused or one of them is deleted. That is, [ws] can be interpreted in three ways:

a) Fusion: 2nd plural [w] (inflection) + 2nd person plural [w] (clitic) –the segment *-s* obviously belongs to the clitic as well.

b) Deletion of the inflectional morph: [ws] stands only for the clitic, the inflectional morpheme is not expressed phonologically.

c) Deletion of the clitic morph: the glide encodes verbal inflection, while *-s* is the sole exponent of the clitic.

Which of these possible analyses is more plausible will be discussed in section 6.2.

- *Locative and Neuter*

Locative and neuter clitics can be analyzed together because they display exactly the same behavior. They have two variants which are phonologically conditioned:

|  | Attached to a vowel | Not attached to a vowel |
|---|---|---|
| **Locative** | [j] | [i] |
| **Neuter** | [w] | [u] |

As shown above, when there is a vowel, either before or after the clitic, the clitic surfaces as a diphthong, either falling (as a coda) or rising (as an onset):

(50)    a.  No  [**j**]  vai
            NEG LOC go.1S
            'I don't go there'

    b.  [**w**]  hem    de tapar
       NEUT have.1PL of cover.INF
       'We have to cover that'

If the clitic cannot be attached to a vowel, we find the vocalic form instead of the glide:

(51)    a.  Comprant- [**u**]
        buy.GER    NEUT
        'Buying this'

       b.  [**i**]  torna    a haver   pujada
          LOC go.back.3S at have.INF increase.S.FEM
          'And there is increase again'

## 2.2.2 Clitic Clusters

All the examples above show the contextual variants in isolation. In combinations, they exhibit different characteristics, as described below:

1. When two clitics are put together, the second clitic appears in its non-vocalic form, as in (52):

   (52)   Això     [**sen**]     diu   mitjons
         this.NEUT REFL+PART say.3S sock.MASC.PL
         'The name for this is socks'

   Whether preceded or followed by a consonant or vowel, the final output is always the same. The reverse (i.e., the full form in the second clitic and the non-vocalic clitic in first position) would be a possible option only if preceded by a vowel, but in actual practice this never happens. A sequence like (53) is impossible:

   (53)   *Això'**s ne** diu mitjons

   Phonotactics should not ban these forms, but they actually never occur. This is discussed in section 6.1.

2. Clitic clusters differ in their syllabification depending on the context. When they are followed by a consonant, as in (52), the second clitic is in coda position. If they are followed by a vowel, the second clitic becomes the onset of that vowel, as in (54):

(54)    [**se  m**]'    arrupís
        REFL DAT.1S shrink.3S
        'It shrinks'

A seen in example (52), the apocopated clitic is chosen when it can be attached to the first clitic vowel. But the same variability found in contractions involving single clitics is also present in clusters. That is, when *se*, *me* and *te* combine with each other (whenever that is possible) or with *ne*, they are not always contracted, as we see in example (55):[24]

(55)    [**se  te**]    moriva'    l        videll
        refl.3S DAT.2S die.3S.IMPF the.MASC.S calf.MASC.S
        'Your calves died'

Again, both variants are used by speakers, although contracted froms are more common. Of the contractable forms, *ne* is the only one which seems to be contracted mandatorily —although it can appear non-contracted in isolation. This is true of all the clitics that it can be attached to, as in combinations with 1st and 2nd person plural clitics, for example:

(56)    a.    Anà- [**βo  n**]    d' aquí
            go.INF INH.2PL PART of here
            'To leave (you) a place'

        b.    [**mo  n**]    comprem dos
            DAT.1PL PART buy.1PL two
            'We buy two of those for ourselves'

In these situations, the contact between the nasal segment of the partitive clitic and the sibilant of the 1st and 2nd person clitics causes the deletion of the sibilant.[25] *[mozne]/*[bozne] would be possible options permitted by the phonology of Catalan, but in fact are never found.

The dative plural can also be combined with the partitive. [lezin] is its phonological outcome:

(57)    [**lezi  n**]    vull    donar  tres, de roses
        DAT.3PL PART want.1S give.INF three of rose.FEM.PL
        'I want to give them three roses'

---

[24]*Me* and *te* yield sequences like [men] and [ten] in combination with *ne*, following the structure of [sen] in example (52).

[25]Although *-s* usually encodes plurality, its deletion in these contexts does not imply a loss of information, as the morphs *mo* and *vo* encode plurality as well.

If a vowel precedes the sequence, the dative gets attached to the vowel, and is thus contracted:

(58)  Dóna-  [**lzi**     **n**]    tres,  de  roses!
      give.IMP DAT.3PL PART three of rose.FEM.PL
      'Give them three roses!'

Although the simple form *les* is preferred for the dative plural, here the marker *-i* always appears in the sequence. [lezne], with the full form of the partitive, would be completely licit according to the phonotactics of Catalan, but it is (almost) never found in Pallarese.[26] It seems that the partitive clitic prefers to be attached to a preceding vowel, and the dative marker *-i* can accomplish this requirement. There is a clear parallel with the non-possible forms *[mozne]/*[bozne], but in this case the strategy to avoid this output is different. The sibilant could be deleted to generate *[len] (as [mon]/[bon]), but instead we find [lezin].

The last possible combination, *li+ne*, does not present any particularity. The partitive appears again in its apocopated form due to the preceding vowel of the dative, which acts as a host for the partitive:

(59)  [**li**     **n**]    dono   tres,  de  roses
      DAT.3S PART give.1S three of rose.FEM.PL
      'I will give her/him three roses'

3. When the 3rd person singular accusative clitic appears in combinations, it has the same behavior as in isolated forms. That is, with the preceding vowel of the first clitic (in accordance with the order described in the previous section, the 3rd person accusative clitic will always be in second position, except when combined with the dative singular or plural), the masculine will surface as [l] but the feminine will keep the gender marker:[27]

(60)  a.  Aguell      enciam       no  [**te**    **l**]       pots
          that.MASC.S lettuce.MASC.S NEG DAT.2S ACC.MASC.S can.2S
          dixar
          leave.INF
          'You can't leave that lettuce'

      b.  La         galleta      [**me**    **la**]   porta      després
          the.FEM.S bucket.FEM.S DAT.1S ACC.FEMC.S bring.3S later

---

[26] One speaker produced these forms, but most participants preferred [lezin].
[27] The first clitic of the cluster can be *me*, *te*, *se*, *mos* or *vos*.

'S/he'll bring the bucket later'

There is no difference in enclitic position, and syllabification follows the same
pattern as with the other clitics that appear in combinations, namely as onsets
if followed by a vowel and as codas before a consonant.

With 1st and 2nd person plural clitics there is also deletion of the sibilant, as
in (56-a) and (56-b). Again, this implies that there is a vowel to which the
clitic can be attached, so the apocopated form for the masculine is favored:

(61)  a.  Aguell      enciam,      porta-       [**mo**
          that.MASC.S lettuce.MASC.S bring.2A.IMPR DAT.1PL
          **l**]!
          ACC.3S.MASC
          'Bring us that lettuce!'
      b.  Mengeu-    [**βo**    **l**],           lo           lluç!
          eat.2S.IMPR INH.2PL ACC.3S.MASC the.MASC.S hake.MASC.S
          'Eat the hake!'

The masculine accusative clitic will therefore never surface as *lo*, and will be
[l] in all circumstances.

With the feminine clitic, the situation is different. The sibilant is elided, but
the gender marker is only lost before a vowel (as when the clitic occurs in
isolation and in (62-b)):

(62)  a.  Mengeu-    [**βo**    **la**],      la          tomata!
          eat.2S.IMPR INH.2PL ACC.3S.FEM the.FEM.S tomato.FEM.S
          'Eat the tomato!'
      b.  La       tomata       [**βo**    **l**]'      amaniu
          the.FEM.S tomato.FEM.S INH.2PL ACC.3S.FEM dress.2PL
          'You dress the tomato'

4. Combinations of 3rd person plural accusative clitics pattern phonologically
   just as they do in isolation. The accusative clitic will surface as *-ls* since it can
   be syllabified with the preceding vowel of the first clitic:

(63)  Les        enciams      [**me**    **ls**]       porta   cada
      The.MASC.PL lettuce.MASC.PL DAT.1S ACC.3PL.FEM bring.2S every
      dia
      day.MASC.S
      'S/he brings me the lettuces everyday'

When the allomorph is *es*, the vowel of the accusative clitic is fused with the vowel of the preceding clitic (reflexive, 1st and 2nd person), so it ends up being -[s] phonetically (as it did in isolation before any vowel):

(64)     Aguetes     tomates     ja     [**te**    **s**]     pots
        This.FEM.PL tomato.FEM.PL already REFL.2S ACC.3PL.FEM can.2S
        menjar
        eat.INF
        'You can already eat these tomatoes'

When *les* or *es*, whether masculine or feminine, are joined to 1st and 2nd person plural clitics, two things can happen:

- In the presence of a liquid, there is deletion of the sibilant belonging to the first clitic:

  (65)    a.   [**mo**    **ls**]     *comprem*
             *dat.1pl* ACC.2PL.MASC buy.1PL
             'We buy these things for ourselves'

        b.   Mengeu-     [**βo**    **ls**]!
             eat.2PL.IMPR INH.2PL ACC.3PL.FEM
             'Eat them!'

  As the sibilant is lost, the accusative clitic can be attached to the vowel, so the contracted form is preferred.

- When the allomorph *es* is attached to *mos* or *vos*, the vowel of the accusative dos not surface and the plural morph /z/ is either deleted or fused with the sibilant of the 1st and 2nd person clitics:

  (66)    a.   Les     tomates     no [**mo**    **z**]
             the.FEM.PL tomato.FEM.PL NEG DAT.1PL ACC.3PL.FEM
             vol     comprar dingú
             want.3S buy.INF nobody
             'Nobody wants to buy us the tomatoes'

        b.   Les     lluços,     [**bo s**]   porten
             the.MASC.PL hake.MASC.PL DAT.2PL ACC.3PL.MASC
             al      restaurant
             bring.3PL to+the.MASC.S restairant.MASC.S
             'They bring the hakes to the restaurant for you'

Thus, in these constructions, *mos* and *vos* stand for *mos + es* and *vos + es*. Whether what is happening is deletion of one of the morphs or instead fusion of the two will also be discussed in section 6.2.

5. When 3rd person accusative clitics appear with dative clitics, some changes occur. In singular, the usual order is reversed, and the accusative clitic precedes the dative (which, in this case, is expressed only by *-i*). Phonologically, they do not behave differently, and the feminine exponent is deleted before a vowel (i.e., the dative *-i*):[28]

(67)  a.  La          tomata       [**l**      **i**]     porta'
          The.FEM.S tomato.FEM.S ACC.S.FEM DAT.3S bring.3S
          l          dilluns
          the.MASC.S monday.MASC.S
          'S/he delivers the tomatoes to him on Monday'

     b.  L'          enciam       [**l**      **i**]     porta'
          The.MASC.S lettuce.MASC.S ACC.S.FEM DAT.3S bring.3S
          l          dilluns
          the.MASC.S monday.MASC.S
          'S/he delivers the lettuces to him on Monday'

If the dative is plural, one might think that the accusative clitic is deleted:

(68)  a.  La         tomata       [**lezi**]  porto'   l
          theFEM.S tomato.FEM.S DAT.3PL bring.1S the.MASC.S
          dimarts
          tuesday.MASC.S
          'I bring them the tomato on Tuesday'

     b.  Lo         lluç         [**lezi**]  porto'   l
          theMASC.S hake.MASC.S DAT.3PL bring.1S the.MASC.S
          dimarts
          tuesday.MASC.S
          'I bring them the hakes on Tuesday'

However, this cannot be the case because in the sequence 3rd person accusative plural + 3rd person dative singular, the output is the same, and the accusative must be present in the combination (the dative is singular and the cluster is plural):

(69)  a.  Les         tomates       [**lezi**]               porto'
          theFEM.PL tomato.FEM.PL ACC.3PL.FEM+DAT.3S bring.1S
          l          dimarts,     amb ell
          the.MASC.S tuesday.MASC.S with he
          'I bring him the tomatoes on Tuesday'

---

[28]As mentioned above, these outputs are opaque and thus it is not clear that the sequence consists of accusative + dative. See section 5.1.3 for more details.

When the combination is preceded by a vowel, contraction occurs, as expected:

(70)    Les        tomates,     diu    que' **[lzi]**
          the.FEM.PL tomato.FEM.PL say.3S that ACC.3PL.FEM+DAT.3S
          portes    lo        dimarts
          bring.2S the.MASC.S Tuesday.MASC.S
          'S/he says that you should bring her/him the tomatoes on Tuesday'

Given the data in (68) and (69), it is clear that the phonological realization of dative and accusative clusters is obscure, something that also happens when both the accusative and the dative are plural, as in (71):

(71)    Les        lluços       **[lezi]**            entrego
          the.MASC.PL hake.MASC.PL ACC.3PL.FEM+DAT.3PL deliver.1S
          lo        dimarts
          the.MASC.S Tuesday.MASC.S
          'I deliver the hakes to them on Tuesday'

In this case, it is not clear whether *les* refers to the dative plural or the accusative plural. Both things are possible since they have the same phonolgical form, so there are two alternative interpretations. On the one hand, one could argue that [lezi] is just a dative plural because the accusative has been deleted. On the other hand, [les] could resort to the accusative plural whereas the dative would be expressed via [i]. An analysis of this puzzle will be offered in section 5.1.3.

As seen in section 2.1, there are certain syntactic restrictions on the combinations with the dative. 1st and 2nd person clitics cannot form sequences with *li*. Thus, the last possible combination targets the neuter clitic *ho*, which gives rise to the output *l'hi*. It seems in this case that *ho* changes its shape and becomes [l] (3rd person singular masculine accusative), whereas the dative is expressed by [i] (cf. Bonet 2002; Mascaró 1986; see Martín 2012 for a different interpretation):[29]

(72)    a.    Demostrant- **[l    i]**,    s'    ho    creurà
               prove.GER    NEUT DAT.3S, REFL NEUT believe.3S.FUT
               'If you prove this to her/him, s/he will believe it'

The combination of the neuter and the dative plural also yields, as in the

---

[29]As previously shown in examples (21) and (23), [laj], and not [li], is what we find in some speakers.

dative plural + 3rd person accusative, the sequence [lezi]/[lzi] (depending on the phonological context), depicted in (73):

(73)    a.    Això,       als            teus            fills,
                  this.NEUT to+the.MASC.PL your.MASC.PL kid.MASC.PL
                  **[lezi]**          dono
                  NEUT+DAT.3PL give.1S
                  'I give this to your kids'

           b.    Això,       als            teus            fills,
                  this.NEUT to+the.MASC.PL your.MASC.PL kid.MASC.PL
                  dóna'          **[lzi]**!
                  give.IMPR.2S NEUT+DAT.3PL
                  'Give this to your kids!'

6. The phonological behavior of locative and neuter clitics clusters is exactly the same as in isolation. If there is a vowel in the preceding clitic, they appear as glides:

(74)    a.    **[te**     **w]**    diu
                  DAT.2S NEUT say.3S
                  'S/he tells you that'

           b.    **[me**   **j]**    porta
                  ACC.1S LOC bring.3S
                  'S/he takes me there'

The full forms of 1st/2nd person singular and 3rd person reflexive clitics show up even when there is a vowel after the cluster, in which case the locative or neuter clitics are syllabified with the following word:

(75)    No **[te**   **j]**   estàs a  casa
          NEG INH.2S LOC be.2S at home.FEM/S
          'You don't stay at home'

But if there is a vowel after the combination, the first clitic can appear in its contracted variant (though this is not very common, so this might reflect the influence of Standard Catalan):

(76)    **[s**     **j]**   amagaran
          INH.3PL LOC hide.FUT.3PL
          'They will hide there'

In these combinations, *ne* is always apocopated, even when the sequence is

followed by a consonant:[30]

(77)  [**n**  **i**]  queden  tres
      PART LOC  stay.3PL  three
      'There are three remaining'

With 1st and 2nd plural clitics, the full forms *mos* and *vos* are chosen (no *-s* deletion):

(78)  a.  [**boz**  **u**]  diu
          DAT.2PL NEUT  say.3S
          'S/he tells you that'

      b.  [**moz i**]  quedarem
          INH.1PL LOC     stay.1PL.FUT
          'We will stay there'

The only problem arises when there is a vowel before the 2nd person clitic. As previously mentioned, the clitic changes its shape to *us* in isolation, and the same holds for combinations:

(79)  a.  No  [**wz**  **u**]  diu
          NEG DAT.2PL NEUT  say.3S
          'S/he doesn't tell you that'

      b.  Emporte           [**wz**  **u**]!
          take.away.2PL.IMPR INH.2PL NEUT
          'Take that away!'

These are, then, the basic forms that 1st and 2nd person plural clitics can take when they combine with other clitics:

| Before *-n, -l* (in combinations) | Elsewhere |
|:---:|:---:|
| mo | mos |
| vo | vos/us (after a vowel) |

In combinations with the locative, the 3rd person masculine accusative clitic surfaces as [li], as in (80):

(80)  Lo        lluç,     a la      nevera,    no
      the.MASC.S hake.MASC.S to the.FEM.S fridge.NEG FEM.S
      [**l**      **i**]  posaré
      ACC.S.MASC LOC  put.FUT.1S

---

[30]See section 6.1.1 for a possible interpretation of these cases.

'I won't put the hake in the fridge'

If the accusative clitic is feminine, it can surface as either [laj] (where the feminine inflectional ending of the accusative acts as a host for the locative clitic, which surfaces as a glide), shown in (81-a), or [li] —(81-b), where the feminine marker is absent and, thus, its shape is the same as in the accusative masculine + locative combination, depicted previously in (80):

(81)  a.  La          tomata,      a la        nevera,    no
          the.FEM.S tomato.FEM.S to the.FEM.S fridge.NEG FEM.S
          [**la**        **j**]  posaré
          ACC.S.FEM LOC put.FUT.1S
          'I won't put the tomato in the fridge'

      b.  La          tomata,      a la        nevera,    no
          the.FEM.S tomato.FEM.S to the.FEM.S fridge.NEG FEM.S
          [**l**          **i**]  posaré
          ACC.S.FEM LOC put.FUT.1S
          'I won't put the tomato in the fridge'

Dative (singular or plural) + locative combinations, i.e., *li/les (hi) + hi*, were excluded from the questionaire because they are not easily elicited. These combinations are not very frequent and surface exactly as the dative in isolation, i.e., [li] (singular) or [lezi] (plural; [les] is more common in isolation, though).[31]

## 2.3   The Definite Article

Definite articles share their form with 3rd person accusative clitics (in both singular and plural), and their different variants are also used in the same phonological contexts. Therefore, the distribution of the definite article in Pallarese Catalan, which mirrors pronominal clitics, is the following:

|                        | *Before a vowel* | *Before a consonant* | *After a vowel* |
|------------------------|:----------------:|:--------------------:|:---------------:|
| **Feminine singular**  | [l]              | [la]                 | [la]            |
| **Masculine singular** | [l]              | [lo]                 | [l]             |
| **Plural**             | [les]∼[es]       | [les]∼[es]           | [ls]∼[s]        |

Table 2.2: *Pallarese Catalan definite articles*

---

[31]The vowel -[i] expresses deixis in both the dative and the locative. See 5.1.1 for more details.

In the feminine, the non-vocalic form of the article appears only before a vowel, as in (82-c). In any other phonological context it surfaces as [la]:

(82)    a.    Les      agafo  **[la]**    mà
                  DAT.3PL take.1S the.FEM.S hand.FEM.S
                  'I take their hand'

          b.    **[la]**      casa
                  the.FEM.S house.FEM.S
                  'The house'

          c.    **[l]**'      espatlla
                  the.FEM.S shoulder.FEM.S
                  'The shoulder'

As was the case with the pronominal clitic, in the masculine [lo] is restricted to preconsonantal environments (if the clitic follows a consonant-final word or it is in absolute initial position), as in (83-a). If there is a previous vocalic host, the article is attached to it —(83-b)—, as well as if the noun starts with a vowel —(83-c)—:

(83)    a.    **[lo]**      llom
                  the.MASC.S loin.MASC.S
                  'The loin'

          b.    Treballa **[l]**      camp
                  work.3S the.MASC.S field.MASC.S
                  'S/he works in agriculture'

          c.    **[l]**      enciam
                  the.MASC.S lettuce.MASC.S
                  'The lettuce'

As for the plural, the gender distinction also gets neutralized in the definite article. The same article is thus used for both the masuline and the feminine plural, but it has three phonological forms (in free variation, see 5.1.4). Before a vowel –(84)– or a consonant –(85)– the article surfaces as either [les] or [es]:

(84)    a.    **[les]**/**[es]** estacions
                  the.PL     season.FEM.PL
                  'The seasons'

          b.    **[les]**/**[es]** enciams
                  the.PL     lettuce.MASC.PL
                  'The lettuces'

(85)    a.    **[les]**/**[es]** tomates
                  the.PL     tomato.FEM.PL
                  'The tomatoes'

b. [**les**]/[**es**] lluços
the.PL    hake.MASC.PL
'The hakes'

When the definite article is preceded by a vocalic host, instead, and it is either followed by either a consonant or a vowel, it can appear as [ls] or [s], as in (86):[32]

(86)    a.    A veure    [**s**]    gavatxos    d' Àreu
to see.INF the.PL French.M.PL of Àreu.M.PL
'To go see the Frenchies of Àreu'

b.    I    [**ls**]    dones    quedaven    aquí
and the.PL woman.F.PL stay.IMPF.3PL here
'And the women stayed here'

All in all, the data show an exact correspondence between pronominal clitics and definite articles. In chapter 5 I will mainly use pronominal data for the theoretical analysis, but as we have just seen, any conclusion drawn for one category is also valid for the other.

---

[32]*Gavatxo* is a pejorative term used to refer to French people.

# THREE

## NOMINAL INFLECTION IN CATALAN AND SPANISH

## 3.1 Gender

This dissertation is about (albeit not exclusively) the exponence of gender, one of the components of the so-called *Phi*-features, which include also person and number. Gender is probably the most salient feature of Romance inflectional systems. Several proposals have been made since the early days of Generative linguistics to analyze it as a formal property of languages from a semantic, morphosyntactic and/or phonological point of view. Although there are correlations between gender and noun inflection based on sex and other semantic criteria (animate vs. inanimate, for example), I will focus instead on the 'formal assignment' of gender (Corbett, 2006), i.e, the arbitrary association of grammatical gender to nouns and adjectives.[1] Specifically, I will try to elucidate the mechanisms that connect the morphosyntactic information of the overall nominal system to the phonological module.

Any theoretical proposal that analyzes gender at the interface between phonology and morphology must answer the following questions:

1. Are inflectional endings exponents of gender or not (i.e., are they simply class markers)?

2. How are inflectional markers stored?

3. What are the mechanisms that turn morphosyntactic information into phonological material?

---

[1]See 4.2.2 for a review of the current proposals that analyze gender from a syntactic perspective and the assumptions that I take from them for this work.

In order to highlight the differences and/or resemblances adopted in my proposal, I summarize in this chapter the most satisfactory analyses of Catalan and Spanish inflectional systems that have tried to answer the above mentioned questions.

### 3.1.1 Catalan

Catalan nouns and adjectives can be decomposed using the following basic template (excluding derivative suffixes attached to the root and number):

(1)     Root + inflectional ending

The template consists of a root that contains lexical information and an inflectional ending at the right edge of the word that has traditionally been considered the expression of syntactic gender. All nouns and adjectives, thus, must reflect this structure in their phonetic outputs, as shown in the following examples:

(2)     cas-[ə]   'house'   (f.)
        llor-[u]   'parrot'   (m.)

Catalan, like most Romance languages, has a double distinction for gender: masculine vs. feminine. The most common inflectional endings for the two genders are depicted in table 3.1:

|  |  | Masculine | Feminine |
|---|---|---|---|
| Inflectional marker | [Ø] | *cor*[Ø]   'heart' | *sal*[Ø]   'salt' |
| Inflectional marker | [u] | *llor*[u]   'parrot' | *mot*[u]   'motorcycle' |
| Inflectional marker | [ə] | *par*[ə]   'father' | *cas*[ə]   'house' |

Table 3.1: *Central Catalan nominal system*

In both masculine and feminine nouns, -[Ø], -[u] and -[ə] are possible endings. However, their distribution is not homogeneous. Due to historical apocope, -Ø is the default class marker for masculine, -[ə] being its feminine counterpart (as in other Romance languages). Thus, the vast majority of masculine nouns do not have an overt morphological marker, as opposed to feminines, which usually do. The other markers can still be considered part of what constitutes the 'regularity' of inflectional exponence (in some cases more than in others), but nevertheless their appearance is by far less common, as indicated by the arrows in table 3.2 (precedence indicates greater frequency). Therefore, we can establish a markedness hierarchy with regard to the exponents found in table 3.1 which varies for each gender (as frequency decreases, markedness increases):

|           | Unmarked | Marked |
|-----------|----------|--------|
| Masculine | Ø        | u>ə    |
| Feminine  | ə        | Ø      |

Table 3.2: *Central Catalan nominal system (unmarked vs. marked)*

Central Catalan has a reduced vowel system in unstressed positions (vowels /o/ and /ɔ/ become [u] in unstressed syllables, whereas /a/, /ɛ/ and /e/ become [ə]), and this is why we only find three different inflectional endings in tables 3.1 and 3.2. We will see later on (in 4.2, when we analyze the nominal system of Pallarese) that the situation differs a bit in Northwestern Catalan (the dialectal group Pallarese belongs to) due to a different pattern of vowel reduction. Most studies regarding Catalan nominal inflection, though, have been devoted to Central Catalan, and this is the reason why I will only concentrate on this variety for the moment.

The markedness hierarchy in table 3.2 is especially relevant to understand the behavior of sibilant-ending nouns like *gos* 'dog', which have a -[Ø] ending in the singular but make use of the marked masculine morph -[u] in the plural, i.e., *goss*[u]*s* 'dogs'. This morphologically conditioned epenthesis, which serves to block contact between the root sibilant and the plural morph -/z/ (*\*goss*), has been a puzzle in morphological studies of Catalan, and several proposals have been made in the generative literature to solve it. This thesis also contributes to a better understanding of this process.

In table 3.2 we can see the most frequent suffixes in the nominal system of Catalan, but these are not the only ones, as the following inflectional markers can be found as well: *-u* (*trib-u*), *-i* (*bigot-i*), *-s* (*temp-s*), *-is* (*cut-is*), *-us* (*vir-us*), *-es* (*càri-es*), *-os* (*cosm-os*). Nevertheless, these marginal endings (which can appear in both masculine and feminine nouns, except for *-os*, which is only found in the masculine) are very peripheral within the system, and are obviously more marked than the exponents shown in table 3.2.[2]

This is, *grosso modo*, the inflectional system of Catalan. It is important to keep in mind that the differences between unmarked and marked endings will become crucial for the analysis presented in chapter 4. Throughout this dissertation I aim to explain the behavior of default nominal elements by a process of lexical insertion framed within Autosegmental Phonology (Goldsmith, 1976). However, this general process of inflection works *only* (and this is crucial) with default nouns, i.e., those that have a -[Ø] ending in the masculine and an -[ə] ending in the feminine. The

---

[2]All nominal classes are analyzed in detail in 7.1. See also Lloret (2013) for a complete review of the inflectional system of Catalan.

other elements are assummed to be lexicalized and, as such, are not produced by the generative device but simply stored in the memory with every lexical entry (this distinction between predictable and lexically specified morphs is assumed in almost all analyses). As noted, we will see in chapter 4 the specific mechanisms that make sure that the overall nominal system works as expected, but I will first briefly review some of the previous analyses regarding Catalan nominal inflection.

### 3.1.1.1 Wheeler (1979)

One of the first attempts to analyze the Catalan nominal inflectional system from a generative perspective is Wheeler (1979). Wheeler assumes for Catalan the generalizations that we have just seen:

a) Feminine words tend to end in -[ə]

b) Masculine words tend to end in -[Ø]

Wheeler's analysis is framed within the SPE model (Chomsky & Halle, 1968), which is rule-based. Besides these default endings, he discusses other markers that are problematic due to different reasons and whose functioning is important to correctly understand the analysis presented in the following sections. On the one side, a rule of right-edge epenthesis is posited for a set of masculine nouns which bear in their roots coda clusters not permitted in Catalan, like stop+liquid (e.g., *centr**e*** 'center' (**centr**).[3] For a different set of nouns, however, this same vowel cannot be considered epenthetic (roots do not contain unsyllabifiable codas) and it must thus be considered part of the stem, as in *infam*[ə] 'infamous' (cf. *blasfem* 'blasphemous') or *mit*[ə] 'myth' (cf. *mot* 'word').

On the other side, a similar scenario is depicted for the inflectional ending -[u]. This vowel needs to be considered lexical in words such as *sur*[u] 'cork' (cf. *mur* 'wall') or *mac*[u] 'nice' (cf. *moc* 'mucus') and, as we saw in *infame* or *mite*, these words require a morphological specification in order to obtain the correct inflectional ending —since otherwise the default rule for inflection in the masculine would make them surface with a -*Ø* ending. Wheeler points out that, by contrast, words ending in a sibilant or a sibilant followed by a stop insert -[u] before the plural morph -*s* (*goss*[us] 'dogs', cf. *gos*). Though the insertion of this vowel is fully productive and very common in Catalan, Wheeler's account stipulates a rule for the insertion of -[u] under certain phonological conditionings, and thus fails to

---

[3]Schwa is the default epenthetic vowel for Central Catalan, whereas -[e] is used in Northwestern Catalan. See 4.1.2 for a description of the epenthetic processes of Catalan.

capture the morphological nature of this 'epenthetic' vowel given that -[u] is the second most common inflectional ending for the masculine after -[Ø], as we saw in table 3.2. The fact that a masculine morph is used in these epenthetic contexts cannot be accidental. That is, Wheeler's account cannot explain why the quality of the epenthesized vowel is different in *centre* compared to *gossos* but it clearly distinguishes between the two epenthetic patterns for masculine nouns, one due to a phonological process and the other lexically specified.

Finally, Wheeler points out that other peripheral markers, such as *-i*, cannot always be treated as suffixes —which he claims to be the case in *bigoti* 'moustache'— , since clearly belong to the stem in words like *canvi* 'change'. We will see also in 7.1.1 how these differences are dealt with in my proposal.

### 3.1.1.2 Mascaró (1985) and Mascaró (1986)

A unitary classification of the nominal system of Catalan and a representational proposal is also provided by Mascaró (1985). The author assumes three different types of inflectional markers:

- Unmarked: -[ə] (feminine) and -[Ø] (masculine)

- Marked: -[Ø] (feminine) and -[u] (masculine)

- Exceptional class: -[ə] or -[i] for the masculine, -[i] or -[u] for the feminine. Other class markers as well, such as -[s] in *temps*, -[is] in *bilis*, etc.

Gender assignment depends on the nature of the underlying representation. For unmarked cases, there is a set of "Gender spelling" rules, shown in (3), that generate default nouns like feminine *casa* 'house' —(3-a)— or masculine *tap* 'cork' —(3-b)— (Mascaró, 1985, pp. 165–166):

(3)    a.    $\text{Ø} \rightarrow [\text{V}] \ /[[\text{X}]_{+N,+F}]$
       b.    $\text{Ø} \rightarrow [\text{E}] \ /[[\text{X}]_{+N,-F}]$

For feminine cases, the gender spelling rule merely provides a V-slot, with no specifications for place features.[4] Then, a "universal convention" (in Mascaró's terms) assigns a schwa to this V-slot, thus generating -[ə] feminine nouns. By contrast, the masculine gender spelling rule provides a -[Ø] ending (i.e., [E]). Whereas the lexical

---

[4]This is exactly what I argue for in this thesis, although the difference is that I assume this process to be general, not restricted only to feminine nouns. That is, under my analysis (see 4), all sntactic functional categories (not only nouns) project a theme position, and this position is always spelled out as V. This captures the idea of the theme vowel in Romance.

entries for unmarked cases consist of bare roots, marked -[u] masculine nouns have a V-slot attached to their representation, as in the adjective *fondo* 'deep' in (4):

(4)     [[fónd][V]]$_{-F}$

An additional gender spelling rule assigns -[u] to this V-position in masculine words (in the context of [V]$_{-F}$), and therefore we can correctly generate masculine-feminine pairs like *fondo/fonda* 'deep' (cf. the lexical entry for the feminine, [fónd]$_{+F}$, which contains only a bare root). As for marked -[Ø] feminines, like *la final* 'the final', the root is also diacritically marked, this time with an empty consituent ([[finál][E]]$_{+F}$). Finally, no rules are needed for the exceptional class, as the vowel is fully specified in the lexical entry.[5]

This proposal makes it possible for the author to capture the behavior of 'epenthetic' -[u] in masculine plurals with unmarked sibilant-ending roots (the case of *gos-gos*[u]*s*, 'dog-dogs'), because the rule that inserts -*u* in the singular also provides the same vowel in the plural in unexpected environments (V-insertion applies freely between sibilants, p. 171). Therefore, it is a step forward with regard to Wheeler (1979) as it links -[u] epenthesis directly to lexically specified -[u] nouns or adjectives like *mac*[u] 'nice'.

Mascaró (1986), which is mainly a descriptive work, treats in depth the nominal system of Catalan. This is the distribution given for the inflectional markers in it, which basically reproduces the assumptions found in Mascaró (1985):

|      | Ø | u | ə | i, is, us, s, etc. |
|------|-----|-----|-----|-----|
| **M.** | *Majority* <br> *dit*[Ø] 'finger' | *Few cases* <br> *sur*[u] 'cork' | *Minority* <br> *di*[ə] 'day' | *Occasional* <br> *bigot*[i] 'moustache' <br> *brind*[is] 'toast' |
| **F.** | *Minority* <br> *sal*[Ø] 'salt' | *Occasional* <br> *mot*[u] 'motorbike | *Majority* <br> *rob*[ə] 'clothes' | *Occasional* <br> *glot*[is] 'glottis', <br> *bil*[is] 'bile' |

Table 3.3: *Inflectional markers (adapted from Mascaró 1986, p.99)*

As we already saw, masculine is realized by -[Ø] and feminine by -[ə]. Although in the masculine -[ə] is more widespread and the -[u] ending is only present in a hundred words (counting only those of frequent use), -[u] must nonetheless be considered the second inflectional marker for the masculine as the strong presence of -[ə] is due

---

[5]Fully specified lexical entries for non-regular endings will turn out to be crucial for my analysis, as well.

to suffixes which already contain it, like *-ista* (*ebenista* 'woodworker'), *-grama* (*pentagrama* 'pentagram'), *-arca* (*monarca* 'monarch'), etc. (and also due to -[ə] endings that have traditionally been considered epenthetic, as in centr[ə] 'center'). For the feminine, although it still constitutes a minority, -[Ø] is quite pervasive (basically because of suffixes like *-ció —concepció* 'conception'; *-or —claror* 'brightness'; *-tud —solitud* 'loneliness'). The remaining inflectional markers are treated as marginal.

Given these facts, Mascaró (1986) adopts a mixed solution for the analysis of Catalan nominal inflection, for which he acknowledges two possible opposite (i.e., 'extreme') solutions:

a) A system where nouns and adjectives would be assigned an inflectional marker randomly. That is, *-Ø*, *-ə* and *-u* could be labelled A, B, and C classes and thus nouns and adjectives would be realized according to the marker they bore (which would be specified in the root). No reference to gender would be needed.[6]

b) A clear-cut gender system, where masculine would always be *-Ø* and feminine *-ə*.

As neither of these options is really true for Catalan, Mascaró (1986) adopts an intermediate position. Gender and class are considered independent but related domains, basically because the feminine implies class B as the unmarked case, class A being the unmarked for the masculine.[7] Therefore, A is by default applied to the masculine and B to the feminine, whereas the other endings are specified as class markers in the root, no matter what gender they have.

Mascaró (1986) also presents the underlying forms of 3rd person pronominal clitics in Barceloní, which will be important for the subsequent analysis of Pallarese:

(5)

|  | Masculine | Feminine |
|---|---|---|
| **Accusative** | l | lə |
|  | lz | ləz |
| **Dative** | li | li |
|  | lzi | lzi |

According to him, /l/ is the 3rd person morph, while gender exponence in the accusative follows the unmarked regular pattern, that is, *-Ø* for the masculine and

---

[6]The term 'class' makes reference to the notion introduced by Harris (1985) whereby gender is completely dissociated from inflectional class. His proposal is reviewed in 3.1.2.

[7]"Aquests dos conjunts són independents, però estan relacionats, bàsicament pel fet que el femení implica com a cas no marcat la classe B i el masculí la classe A." (Mascaró 1986, p. 41)

-ə for the feminine. Gender is not expressed in the dative, which is assumed to be
-*i*. Later on, in 5, we will modify these morphosyntactic assumptions for Pallarese.

### 3.1.1.3 Viaplana (1991)

Viaplana (1991) establishes the following rules for gender and number realization
(rules 20 and 21 in Viaplana 1991, p. 241):[8]

(6)    Rule 20: Gender $\longrightarrow$ ə /[+fem]

(7)    Rule 21: Number $\longrightarrow$ s /[+pl]

Given that not all feminine words take -[ə] and -[s] as feminine and plural endings,
respectively, words that deviate from this pattern are lexically specified as not un-
dergoing the rules in (6) and (7). This is what happens with feminine nouns like
*llum*[Ø] 'light' — (8) — or invariable nouns like *fons* 'bottom' —(9)—, which have
the same shape in singular and plural (*fons profund* 'deep bottom/*fons profunds*
'deep bottoms'):[9]

$$
(8) \quad llum \text{ 'light'} \quad \begin{bmatrix} +\text{N} \\ +\text{fem} \\ \pm\text{pl} \\ -\text{rule 20} \end{bmatrix}
\qquad
(9) \quad fons \text{ 'bottom'} \quad \begin{bmatrix} +\text{N} \\ -\text{fem} \\ \pm\text{pl} \\ -\text{rule 21} \end{bmatrix}
$$

No rule is needed to provide the inflectional marker of regular masculine singular
nouns, as -[Ø] (absence of features) is the default ending (e.g., *cor*[Ø] 'heart'). As
for masculine plurals with a -[u] ending, like *gos-goss*[u]*s* 'dog-dogs', this vowel is
considered purely epenthetic, although morphologically conditioned. The fact that
feminine plurals do not epenthesize at all in the same contexts, i.e, when the root
ends in a strident (or a strident plus a plosive) and the plural morph is attached to
it, as in *pols* 'dust' (*la pols* sg. - *les pols* pl.) leads Viaplana to conclude that the
rule for -[u] epenthesis applies only in [–fem] contexts, as in (10) (Viaplana, 1991, p.
249):

(10)    Rule 43: Ø $\longrightarrow$ u /[+strident]([+plosive])_s##
          (*only applied to* [–fem])

---

[8]Contrary to Mascaró (1986), Viaplana (1991) assumes that the plural morph is underlyingly
voiceless, i.e., /s/. I follow Mascaró (1986) in considering the plural morph to be /z/ for the reasons
set forth in 3.2.

[9]Viaplana (1991, p. 244) gives this morphosyntactic representation for both nouns in singular
and plural, hence the symbol [±].

Feminine nouns like *pols*, instead, undergo a different rule that degeminates two consecutive sibilants, those that belong to the root and the plural morf (pols+s > pols), and the same applies to some masculine nouns that end in a sibilant like (9), which deviate from the common pattern and do not insert -[u] in the plural, i.e., \**fons*[u]*s* (therefore, these nouns must be marked as [– rule 43])

Masculine nouns and adjectives like *magr*[ə] 'lean' are considered to undergo regular epenthesis (\*ma[ɣr]), as in the previous analyses of Wheeler (1979) and Mascaró (1985, 1986). This contrasts with masculine nouns like *poet*[ə], whose final vowel cannot be inserted for phonotactic reasons and is instead considered a genuine feminine marker. Thus, rule 20 in (6) (with a slight modification, p. 252) needs to be added to their lexical entry so that it can provide the vowel -[ə] to a masculine noun.

For nouns like *municip*[i] 'municipality', whose final vowel disappears in derivatives like *municipal* (\**municip**ial*) 'from the municipality', as opposed to *oc*[i] 'leisure' - *ociós* 'at leisure', -[i] is not considered an inflectional element, and an allomorphic solution is adopted instead, i.e., {/munisipi/,/munisip/}.

Some Catalan nouns also bear a -[u] ending, as in *carro* [káru] 'carriage', *monjo* [móɲʒu] 'monk'/*ganxo* [gáɲʃu] 'hook', *borratxo* [burátʃu] 'drunken' or *mono* [mónu] 'monkey', and different interpretations are given for these cases. First, Viaplana claims that a trill cannot occur in word-final position and, thus, there is a process of *morphologically conditioned epenthesis* in nouns like *carro*.[10] The same idea is applied to words like [móɲʒu] or [gáɲʃu], whose final tautomorphemic coda cluster is not tolerated in Catalan, and therefore rule 43 in (10) is extended to these other noun classes.[11] The notion of morphologically conditioned epenthesis, also used in Lloret & Viaplana (1992), is crucial for the analysis developed in this thesis, and in section 7.1.1 it will be shown that cases like *monjo* and *ganxo* can be easily analyzed with the regular process of gender spell-out proposed for (Pallarese) Catalan.

Viaplana argues that, for words like borratxo, there are two alternative explanations, both of them potentially valid. Either the vowel -[u] is part of the stem in the lexical entry, or it is a product of morphologically conditioned epenthesis, which means that the context for the application of rule 43 in (10) must be extended once again. In both cases, these nouns need to be specified somehow to differentiate them from nouns like *esquitx* [əskítʃ] 'splash' which end in a final affricate.

---

[10]Viaplana considers the trill to be, underlyingly, a sequence of two taps.

[11]As Viaplana notes (p. 256), this cluster is licit in plurals like *banys* [báɲʃ] 'bathroom', where the sequence -[ɲʃ] is a product of a progressive assimilation (the plural morph /z/ is palatalized to [ʃ]), but not in *monjo/garxo*-like words, where there is regressive assimilation instead.

Finally, in *mono* [mónu] there is no doubt that the vowel must be lexically specified. Under no circumstances can -[u] be considered epenthetic, and Viaplana adopts an allomorphic solution to account for the distribution of these nouns, which implies that *mono* (and *borratxo*, if -[u] is considered part of the stem as well) has two entries in the lexicon, one restricted to inflection in the singular and plural, i.e., /monu/ for *mon*[u]-*mon*[u]s, and the other for derivation, e.g., /mon/ for *monet* (\**mon**u**et*) 'little monkey'.

To conclude, Viaplana acknowledges that non-linear phonological analyses, albeit not adopted in Viaplana (1991), might be better suited for the intricacies of Catalan nominal morphology, and this is in fact the solution implemented in Lloret & Viaplana (1992), which gives a more unified analysis of different phenomena (especially those regarding masculine nouns and -[u] epenthesis).[12]

### 3.1.1.4 Lloret & Viaplana (1992, 1997); Lloret (1998)

Lloret & Viaplana (1992), which depart from previous works in some respects, provide an autosegmental analysis (Goldsmith, 1976) arguing the following:

1. The vowel -ə can have three different values:

   (a) It can be the regular feminine marker, introduced by a gender rule: ([fem]→[ə]).

   (b) In masculine nouns like *poet*[ə] 'poet' it is part of the root and is thus lexically stored (as opposed to Viaplana 1991, who considers this vowel a feminine ending).

   (c) It can be the product of epenthesis in sequences which cannot be incorporated into the prosodic structure, as in *simpl*[ə] 'simple' (\*sim[pl]).

2. Masculine default nouns have no gender marker (-Ø represents the absence of features). Feminine nouns like *flor* 'flower' need to be lexically marked with a diacritic that specifies the non-application of the gender exponence rule.

3. Masculine -[u] nouns are split into two categories. On the one hand, -[u] is a morphologically conditioned epenthetic segment specified in the lexical entry of certain words whose syllabification would not be possible otherwise (*rotll*[u] 'roll', \**ro*[tʎ]). While -[ə] epenthesis is postlexical, the epenthesis of -*u* is lexical, and it applies also to plurals whose root ending consists of a

---

[12]In fact, Mascaró (1985)'s treatment of nouns with lexical -[u] like *mon*[u], and those undergoing epenthesis like *monjo* (or plurals like *gossos*), is superior to Viaplana (1991)'s, as they are grouped together with a rule that supplies a V-slot in both cases.

strident (or strident + stop sequence), as in *cas* 'case' > *cas*[u]*s* 'cases'. On the other hand, in words witout any syllabification problem, a floating segment *-u* (which is stored with the lexical entry) is linked to the prosodic structure *via* a 'root extension rule' (*regla d'extensió del radical*) that has previously created a skeletal position. The same 'root extension rule' applies to nouns whose word ending is -[i], as in *bigoti* 'moustache'.

4. Words with rare endings like *-s* (*temps* 'time'), *-us* (*cactus* 'cactus') or *-is* (*bilis* 'bilis') are assumed to have a floating segment *-s* which can be linked to the prosodic structure by the special behavior of this segment in the phonology of Catalan.[13]

Lloret & Viaplana (1992) offer many interesting insights, especially those regarding the treatment of the inflectional marker *-u* and its relation to epenthesis from an autosegmental perspective. The 'root extension rule' proposed there, which inserts a skeletal position without specifying any vowel quality, is similar to the idea that will be developed later on in chapter 4 whereby gender features are always realized in a specific morphological position, a proposal that has the advantage of providing a more unified analysis for phenomena that are dealt with separately in Lloret & Viaplana (1992).

Lloret (1998) maintains the essence of Lloret & Viaplana (1992)'s analysis (extended there to Spanish as well), but it introduces some important ideas which will be important to understand Lloret & Viaplana (1997).[14] Using diminutives as evidence for the classification of the nominal system, it determines, following Harris (1992), that there is a 'stem extension rule' (*regla d'extensió del lexema*) which guarantees the insertion of a vocalic slot without any melodic content next to the root. This skeletal position can be filled in three different ways according to the three different patterns of the nominal system:

a) Regular class: gender features (such as -[ə] in the feminine)

b) 'Additional inflectional marker' class (*formatius adicionals*): non-regular inflectional endings which block the default gender assignment (such as *-i* in *bigot*[i] 'moustache')

---

[13]Complex codas in Catalan can have *-s* as the second or even third segment of the coda without respecting the principles of the Sonority Hierarchy. This leads the authors to argue that it has a 'privileged' status.

[14]Although Lloret (1998) is published in 1998, it is based on the analysis presented at the *XXI Congresso Internazionale di Filologia Romanza*, held in Palermo in 1995 (18-24 September). This means that it is prior to Lloret & Viaplana (1997), although it does neet seem so by the publication dates of both articles.

c) 'Unclassified' endings (*desclassat*): special diacritics of the root (as in *map*[ə] 'map')

The basic idea is that the three groups presented above exhibit different behavior regarding diminutive formation. In a), when a diminutive suffix is added to the stem, the regular gender marker is also assigned to it, as illustrated in (11):

(11)     *cas+a* (f.) 'house' > *cas+et+a* 'little house'

In b) the inflectional marker is not regular, but its diminutive inflection follows instead the regular pattern, as illustrated in (12):

(12)     *flor* (f.) 'flower' > *flor+et+a* 'little flower'

The difference arises in class c). When a diminutive is created, as shown below, it inherits the inflectional marker of the stem:

(13)     *Carle-s* (m.) > *Carl-et-**s***

In a later work, Lloret & Viaplana (1997) propose a similar classification for Catalan and Spanish, using word markers along the lines of Harris (1985, 1991, 1992). They propose that *-a* (feminine) or *-o* (masculine) for Spanish and just *-*[ə] (feminine; the masculine has no exponent) for Catalan are directly related to gender. All other words ending in a non-canonical element, however, "are lexically specified as having an idiosyncratic extra post-stem vowel, i.e., Harris' word markers" (Lloret & Viaplana, 1997, p. 177). The application of gender-marking rules is blocked once the word marker has been specified in the lexical entry and, thus, the classification proposed there is a mixed system with some endings being a direct reflection of gender and others not.

One of the article's claims is also that the set of words whose ending does not match the expected gender in diminutives is "lexically marked for triggering the opposite gender-marking rule" (Lloret & Viaplana, 1997, p. 184). That is, for a word like *mapa* 'map' we would expect the diminutive \**mapito* 'little map', as non-regular endings are usually only tolerated in the base forms, but we find *mapita* instead. The fact that *-a* is maintained in the diminutive leads the authors to argue for a diacritic in the root that selects the opposite default gender exponent. We will see in 7.1 how these problematic cases can be dealt with in the analysis proposed in this thesis without resorting to special diacritics.
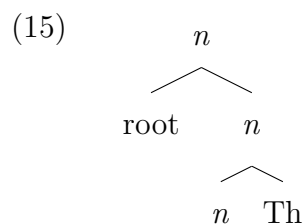
### 3.1.1.5  Oltra-Massuet (1999)

Oltra-Massuet (1999) offers a new formal analysis of Catalan morphology within Distributed Morphology (Halle & Marantz, 1993) and, although it concentrates on verbal inflection, there is a section where the morphosyntactic structure of verbs is extended to nominal categories.[15]

In a framework like Distributed Morphology roots are acategorial, and thus a category defining head needs to be joined to the root to create a noun, as in (14):

(14)

$$
\begin{array}{c}
n \\
\diagup\diagdown \\
\text{root} \quad n
\end{array}
$$
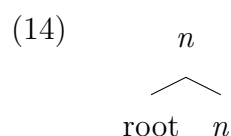
Later on, a well-formedness condition is imposed on syntactic functional heads that requires the adjunction of a postsyntactic theme position:

(15)

$$
\begin{array}{c}
n \\
\diagup\diagdown \\
\text{root} \quad n \\
\diagup\diagdown \\
n \quad \text{Th}
\end{array}
$$

This well-formedness condition which creates a theme position in the overall morphological structure of Catalan is the key innovation provided by Oltra-Massuet (1999). Theme vowels have always been at the core of Romance verbal inflection studies, but Oltra-Massuet extends this idea of an autonomous morphological position to *all* syntactic functional heads. That is, in a verbal form like *cantava* 'S/he sang', there are two theme vowels, one for little *v* (*cant**a**va*), which is the one that has traditionally been considered the theme vowel of Romance verbs, and the other for Tense (*cantav**a***). This idea, as already mentioned, is not restricted to verbs, since nouns, adjectives or determiners also bear this morphological position. This assumption is crucial to properly understand the analysis presented later in chapter 4, as I also assume for Pallarese Catalan the well-formedness condition proposed by Oltra-Massuet (the phonological spell-out for gender in nouns and clitics cannot be understood without the theme position).

Regarding the nominal system, Oltra-Massuet's basic claim is that roots carry class markers which are realized in the theme position after Vocabulary Insertion, thus unifying two apparently unrelated notions, although she does not elaborate

---

[15]Harris (1999) also argues for a theme position in the Spanish nominal system (see 3.1.2), but Oltra's proposal is more fully developed.

further on the topic.[16] However, the use of class markers can be dispensed with without removing the theme position once we have determined the bases for gender spell-out, as will be shown in chapter 4.

All in all, the theme position developed in Oltra-Massuet (1999) is a powerful representational tool that helps us to understand the behavior of nominal inflection and links it to the overall structure of Catalan, thus unifying apparently unrelated phenomena.

### 3.1.1.6 Bonet et al. (2007)

To my knowledge, the newest proposal for the formal analysis of gender in Catalan is Bonet et al. (2007). The radical departure from all previous literature studying Catalan gender exponence is, besides the use of Optimality Theory (Prince & Smolensky, 2004), the idea that the phonological component controls the assignment of exponents.[17] That is, although morphology matches syntactic features with phonological exponents, these exponents are not fixed but ordered instead in a set of allomorphs that relies on markedness (more marked = less frequent, i.e., the most frequent allomorphs will be given preference). The constraint ranking of the phonological module, then, decides in all cases which allomorph will win. The lexical entries for gender in Catalan provided by the authors are shown in (16):

(16)     Masculine:  /{Ø>u>ə}/
         Feminine:   /{ə>Ø}/

The input for gender is, thus, rich. Their claim is that in unmarked cases, i.e., in words that follow the regular inflectional pattern of the language, the ordering relation of the allomorphs (different for each gender) is preserved by the action of the contraint PRIORITY, shown below. PRIORITY guarantees that the first allomorph of the hierarchy surfaces in the output:

(17)    PRIORITY
        Respect lexical priority (ordering) of allomorphs. (Mascaró, 2007)

The input for a masculine word like *pas* 'step' is /pas{Ø>u>ə}/, and PRIORITY favors the output [pás] over *[pásu] or *[pásə] due to the hierarchy established in the input. However, not all nouns in Catalan bear non-canonical inflectional endings

---

[16] "My intention is to demonstrate that our proposal for verbs extends far beyond the scope of verbal morphology, in that it unifies two as yet unrelated notions, to provide a first approximation to a unified analysis of stress." (Oltra-Massuet, 1999, p. 80)

[17] Wheeler (2005, p. 263) also provides an OT analysis for some inflectional phenomena of the nominal system, but not in a systematic way, as he focusses basically on phonological processes.

for masculine or feminine, and marked nominal elements need a subcategorization requirement to avoid surfacing with the default exponent (this diacritic is similar to the morphological specifications of roots found in the previous works that have already been analyzed), as in (18):

(18)     Masculine: /paɾə, /ʎɔɾu/   Feminine: /salø/

The words *sal*[Ø] 'salt', *par*[ə] 'father' and *llor*[u] 'parrot' do not behave as expected according to the regular inflectional process of Catalan, and thus some mechanism must prevent them from taking the masculine/feminine default exponent. If only PRIORITY were active in the language, this default exponent would always surface, and thus another constraint, RESPECT (which must necessarily be ranked higher than PRIORITY), is needed to generate the marked pattern:

(19)    RESPECT
        Respect idiosyncratic lexical specifications (Bonet et al., 2007, p. 918)

The effects of subcategorization and constraint ranking are shown below (examples extracted from Bonet et al. 2007, p.919), for both unmarked —(20)— and marked nouns —(21):

(20)

| /pas-{Ø>u>ə}/ | RESPECT | PRIORITY |
|---|---|---|
| a. ☞ pás | | |
| b.   pásu | | 1W |
| c.   pásə | | 2W |

| /tak-{ə>Ø}/ | RESPECT | PRIORITY |
|---|---|---|
| a. ☞ tákə | | |
| b.   ták | | 1W |

(21)

| /mos$_u$-{Ø>u>ə}/ | RESPECT | PRIORITY |
|---|---|---|
| a. ☞ mósu | | 1 |
| b.   mós | 1W | L |
| c.   mósə | 1W | 2 |

| /sal$_Ø$-{ə>Ø}/ | RESPECT | PRIORITY |
|---|---|---|
| a. ☞ sál | | 1 |
| b.   sálə | 1W | L |

*Pas* 'step' and *taca* 'stain', which follow the unmarked pattern for gender realization, do not have any subcategorization requirement and RESPECT cannot therefore have any effect on their exponence. By contrast, *moss*[u] 'guy' and *sal*[Ø] 'salt' are subcategorized to surface as -*u* masculine and -*Ø* feminine nouns, respectively, and

this is the reason why RESPECT favors these allomorphs. Even if they incur violations of PRIORITY, these candidates are preferred over those with default exponence by the ranking RESPECT >> PRIORITY.

The advantage of this proposal becomes clear when we turn to plural formation. As already noted, masculine plurals of nouns and adjectives ending in a sibilant exhibit a special behavior because they introduce a marked masculine inflectional marker between the root and the plural morph:

(22)    [pás] 'step' > [pásus] (*[páss]) 'steps'

The contact between two sibilants is dispreferred in Catalan. In order to avoid this OCP-sibilant phonological problem, this set of nouns and adjectives resorts to morphology.[18] Recall that inflectional allomorphs are ordered in a markedness hierarchy where the allomorph *-u* stands in second position. The interaction between the constraint PRIORITY and a phonotactic constraint OCP-SIBILANT generates the correct output, which uses input material (an inflectional marker) to repair syllable strcuture. The use of PRIORITY can cope with many of the puzzles of Catalan noun inflection, but later in section 4.2.1 we will see that this analysis is not problem-free and which alternatives can deal with the same inflectional phenomena.

### 3.1.2  Spanish

Like Catalan, Spanish nouns can also be decomposed under the template Root + inflectional ending. The inflectional markers used by the two languages are similar, although their default endings are not the same, as illustrated in (23) for Spanish:

(23)    *cas-a*  'house'  (fem.)
        *lor-o*  'parrot'  (masc.)

Whereas *-a* is shared by the two languages as the default feminine ending, *-o*, and not *-Ø*, is the default masculine ending for Spanish. Besides *-o* and *-a*, also *-e* and *-Ø* are common inflectional markers, as shown in table 3.4:

---

[18]The *Obligatory Contour Principle* (OCP) was first used outside tonal phonology by McCarthy (1986), where it was defined as follows: "at the melodic level, adjacent identical elements are prohibited".

|  |  | Masculine | | Feminine | |
|---|---|---|---|---|---|
| Inflectional marker | [a] | *dí*[a] | 'day' | *cas*[a] | 'house' |
| Inflectional marker | [o] | *lor*[o] | 'parrot' | *man*[o] | 'hand' |
| Inflectional marker | [e] | *hombr*[e] | 'father' | *suert*[e] | 'luck' |
|  |  | *hul*[e] | 'oilcloth' | *clas*[e] | 'class' |
| Inflectional marker | [Ø] | *jamón*[Ø] | 'ham' | *cruz*[Ø] | 'cross' |

Table 3.4: *Spanish nominal system*

As can be seen in table 3.4, all endings can be used for both masculine and feminine nouns, but of course their frequency of appearance varies substantially depending on the gender of the word. In fact, one of the main goals of the morphological studies regarding Spanish nominal inflection is to determine whether there is a direct correlation between inflectional markers and gender (i.e., *-a* = fem, *-o* = masculine). In addition, there has been a long-standing debate about the epenthetic status of the vowel *-e* that appears in certain words, as there is a subdistinction within the *-e* group, as shown in (24):

(24)  a.  *hul*[e], *clas*[e]
      b.  *hombr*[e], *suert*[e]

While in (24-a) the nouns could be phonotactically well-formed without bearing any exponent, as in \**hul* and \**clas*, which end in tolerated codas in Spanish, the same does not hold for (24-b). The codas in \**hom*[br] and \**sue*[ɾt] are illicit in Spanish and some authors have claimed that the appearance of *-e* is due to a process of epenthesis. In fact, epenthesis would apply not only to this set of nouns but also to the plurals of some *-Ø* nouns, as shown in (25):

(25)   *cruz*[Ø]  >  cruc[es]

The fact that plurals like \*cru[θs] are not possible in Spanish by syllable well-formedness requirements would force the insertion of an epenthetic vowel *-e*, according to these authors. Some others, however, argue that this vowel is in fact inflectional. Be that as it may, let us review the most important studies of the Spanish nominal inflection in this section in order to shed light on the above-mentioned puzzles. We will include in our discussion the status of other non-regular inflectional endings such as *-s* (*Carlos* 'Charles'), *-is* (*brindis* 'toast'), *-us* (*virus* 'virus'), etc.

### 3.1.2.1 Harris (1985, 1991, 1992, 1999)

In Spanish, one of the authors who has treated gender most extensively from a morphological and phonological point of view is Harris. Harris (1985) coined the term 'word marker' to refer to inflectional endings, and this term gained popularity for the analysis of inflection not only in Spanish but also in other languages (as we have already seen for Catalan). Word markers "make no syntactic or semantic contribution to the words they are part of; their only grammatical properties are their phonological form and their distribution, which is unlike that of any other class of morphemes" (Harris 1985, p. 34). This move by Harris implies *de facto* that word markers are morphomic in nature (Aronoff, 1994), as their only task is morphological, i.e., they classify nouns, but they are not involved in agreement or any other syntactic function.

The core of the proposal is that word markers are floating morphemes (Autosegmental phonology, Goldsmith 1976), in the sense that roots have their melody linked to the skeleton in the lexicon but word markers do not, i.e., the melody is floating in a separate tier, as in (26):

(26)    *sed*[a] 'silk'

$$/ \quad \text{s} \quad \text{e} \quad \text{d} \quad \text{a} \quad /$$
$$| \quad | \quad |$$
$$\text{C} \quad \text{V} \quad \text{C}$$
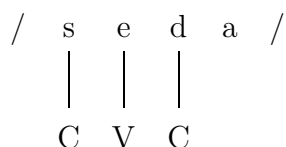
In order to receive phonetic interpretation, a rule labeled 'Slot supply' provides skeleton slots that allow the connection to the prosodic structure.[19] As this rule can only apply at the word level, word markers disappear in suffixed words at the stem level (*sedoso* 'from silk', and not *\*sedaoso*, derived from the noun *sed*[a] 'silk'), because they cannot be linked to any skeletal position.

Although it is argued that nouns are assigned arbitrary word markers, it is obvious to any Spanish speaker that their distribution is asymmetrical. Indeed, Harris proposes the following hierarchy (adapted from Harris 1985, p. 42):

(27)

---

[19]Although the assumptions in my proposal are different in many other respects, we will see in chapter 4 that the insertion of an underspecified theme vowel for all nominal elements is reminiscent of this idea of 'Slot supply'.

| Regular | | Irregular |
|---|---|---|
| *Gender-marked* | *Not gender-marked* | |
| A-1 | A-2 | B |
| | | man*o* |
| hart*o* | | dí*a* |
| ar*o* | part*e* | trib*u* |
| cart*a* | par | obo*e* |
| ar*a* | | síntes*is* |
| | | … |

A-1 consists of what can be regarded as the 'default option' for Spanish, i.e., masculine *-o* nouns and feminine *-a* nouns. A-2 contains *-e* words, as well as those with no vowel ending, regardless of gender. B, by contrast, is a small class, and thus we can regard (B) as marked vs. unmarked (A-1, A-2), A-1 being the unmarked option also within the A class.

In short, the formal proposal implies that A-1 nouns are assigned the default word marker by a lexical redundancy rule ('Marker Spell'), so that no other information is needed in the lexical entry (Marker Spell features can be interpreted phonetically thanks to the insertion of a vocalic position provided by Slot Supply). A-2 nouns, on the other hand, do not contain information about word markers either, but need to have specified in them the non-application of the Marker Spell rule. This generates words like *par* 'pair' (m.), without the default *-o* ending, but also words like *parte* 'part' (f.), where *-a* does not surface but *-e* is instead inserted for syllabification purposes. Finally, class B nouns contain both the morphological feature 'no Marker Spell rule' and the specification of the word marker in their lexical entries.[20] Specifications in lexical entries and the application (or not) of certain morphological and phonological rules determine, thus, the final distribution of the nominal system in Spanish.

Expanding the analysis set out in Harris (1985), Harris (1991) distinguishes three instances of what has traditionally been labeled gender:

1. biological sex

2. grammatical gender (masculine vs. feminine)

---

[20]A word is assigned to class B if *-e* cannot be related to syllable structure, as in *hule* 'oilcloth' or *oboe* 'oboe'. Furthermore, Harris's analysis can be extended to plurals, since the appearance of *-e* in the plurals of *-Ø* words, like *sol* 'sun' > *soles* 'suns', is due to the insertion of a skeletal position which is not phonetically filled in the singular but is still available when the addition of the plural morph provokes a phonotactic problem —*[sols].

3. form class (labelled 'word marker' as well)

According to the author, these are independent but interrelated domains, as "sex is a matter of semantics and/or biology, gender is involved in syntax-dependent concord, and form class is a matter of the morphophonology of individual lexical items." (Harris, 1991, p. 28). Harris acknowledges that *-o* and *-a* are overwhelmingly used to mark masculine and feminine nouns, respectively, but the correlation is not absolute.

Of the three, only grammatical gender plays a role in syntactic agreement, as shown below:

(28)   Un         poet**a**      buen**o**
       INDEF.M.S poet.MASC.S good.M.S
       'A good poet'

Although the noun *poeta* has *-a* as a word marker, its syntactic gender is masculine. Thus, the adjective, which follows the regular pattern of Spanish, agrees with it and is inflected with the default *-o*.

Harris distinguishes the following word markers for nouns, adjectives and adverbs in Spanish:[21]

| Word marker | Example | |
|---|---|---|
| **-o** | *muchacho* | 'boy' |
| | *mano* | 'hand' |
| **-a** | *muchacha* | 'girl' |
| | *día* | 'day' |
| **-e** | *héroe* | 'hero' |
| | *inmune* | 'immune' |
| **-Ø** | *mar* | 'sea' |
| | *mujer* | 'woman' |
| **-i** | *cursi* | 'in bad taste' |
| **-u** | *tribu* | 'tribe' |
| **-Vs** | *lejos* | 'far' |
| **-s** | *quizás* | 'maybe' |

Table 3.5: *Spanish word markers (adapted from Harris 1991)*

---

[21] Adverbs are included in this list because, although they do not bear gender features, they also make use of word markers. An argument in support of this idea is the fact that in derivational processes the word marker is lost, as in *lejos > lejano* ('far' (adv), 'distant' (adj.). *\*lejosano*). This leads him to the conclusion that there should be a distinction between word markers and gender morphemes.

Within this set of word markers, Harris establishes a three-way distinction according to productivity and numerical preponderance (similar to what is proposed in Harris 1985):

| | | |
|---|---|---|
| **Inner core** | *-o* *-a* | Regular |
| **Outer core** | *-e* *-Ø* | Regular |
| **Residue** | Other word markers | Irregular |

Table 3.6: *Regular and irregular word markers according to Harris (1991)*

Thus *-o* (masculine), *-a* (feminine) and *-e* (much less frequent, but present in both genders) would be the regular word markers, whereas the rest would be labeled irregular. It is important to note that, within the residual type, *-a* is relatively common in the the masculine (*profet*a, *telegram*a, etc.), with almost 600 exemplars —as well as quite a few adjectives, whereas *-o* in the feminine is extremely rare (only *man*o is widely used, and not a single adjective bears it.). This quantitative difference has consequences for the morphological machinery used by Harris, since exceptional non-feminine *-a* nouns and adjectives will be derived by the generative engine, while the rest, including non-masculine *-o* lexical items, will be excluded from the morphological system as they are assumed to be lexicalized.[22]

As for the formal analysis of the facts just discussed, Harris, developing his idea about the non-correlation between gender and declension class, splits the nominal system into three dimensions:

- Category (noun, adjective)

- Gender

- Class

Regarding gender, only the feature femenine is present in the lexical entry (in the absence of features, they are given the unmarked masculine value) which is introduced in the derivation by Feminine Marker and Marker Realization rules (the former providing morphological information, the latter phonological), as shown below in (29):[23]

---

[22]This distinction between generative engine and lexicalization will be pushed further in my analysis of Pallarese Catalan, where all non-default inflectional endings are lexicalized. See 7.1 for details.

[23]As Harris notes, these rules provide information that is predictable from *morphological* information, unlike phonological redundancy rules, which provide phonological information from phonological information.

(29)    a.    f → ]a[24]

      b.    To form the $X^0$ level of nouns, adjectives, and adverbs, insert suffixal /a/ if the stem is marked ]a; otherwise, insert suffixal /o/.

The default word marker *-o* is introduced by the Elsewhere Condition (Kiparsky, 1982a), so that only ]a and ]Ø need to be specified in the lexicon. The lexical items belonging to the 'outer core' *-e* are specified as ]Ø, but *-e* is inserted for syllabification purposes throughout the derivation, as in Harris (1985) (thus, the adjective [bérde] 'green' would have a lexical entry /berd/]Ø –*[bérd]–).[25] The set of words with an *-e* marker not needed for syllabification purposes (like *héro*e or *sed*e) have to be lexically specified, as they do not fit into the morphological machinery that generates the regular patterns. The same holds for the 'residue' established in table 3.6.

Although Harris's analysis can deal with the regularity of the nominal system, it fails to give a fully satisfactory analysis of what he calls 'outer core' (class A-2 in Harris 1985). The epenthetic nature of *-e* in this category does not seem very appealing, for reasons presented later on about the location of epenthesis in Romance (see 4.1.2).

Basically, Harris (1992) assumes what had already been established in Harris (1991). The only change is that nouns are labelled in a different way, creating four distinct classes. Class I is the default, with *-o* masculine nouns and no specification needed for gender in the lexical entry (as it is the elsewhere category). Feminine nouns with *-a* endings are grouped in class II, whereas class IIIA incorporates *-Ø* words like *cruz* 'cross' or *carne* 'meat', with *-e* in the latter case being a product of epenthesis (*carn). As for the IIIB group, "the lexical entries of these words specify by brute force that they have an unspecified vowel (manifested phonetically as the default vowel *-e*) in stem-final position" (Harris, 1992, p. 73), as in *hero*-V > *héroe* 'hero'. Finally, in class IV we find *-Vs* and *-s* words like *dosis* 'dose' and *tórax* ([tóraks]).

Harris (1999) changes substantially the views adopted in previous analyses. Framed within Distributed Morphology (Halle & Marantz, 1993), he deals with Spanish word structure by using as a starting point nasal depalatalization, a process in which Spanish palatal nasals (and laterals) are banned in absolute coda position:

(30)    *desdé*n 'scorn' but *desdeñar* 'to scorn'

Although previously analyzed as an argument for cyclicity (the plural is *desdenes*,

---

[24]Square brackets indicate the end of the root.

[25] "(...) when this position is syllabified, phonological redundancy rules fill it with the features of the *maximally unspecified default vowel e*. —emphasis mine—" (Harris, 1991, p. 57)

depalatalization applies in the first cycle before plural formation —Harris 1983), Harris rejects this idea and gives a new classification for the Spanish nominal system that solves the depalatalization problem.

It is assumed there that in the Morphological Component a thematic suffix, i.e., the form class morpheme, is adjoined to the root, something very similar to what is proposed in Oltra-Massuet (1999). Again, Harris argues that the word markers appearing in this position are purely morphological and have no syntactic function whatsoever. Phonological shape determines class membership, something that appears as well in adverbs (which, obviously, do not bear gender features). With the assumptions of previous analyses (Harris 1985, Harris 1991 and Harris 1992) the Vocabulary Insertion rules for the thematic suffix position should be the following:

(31)  III  $\leftrightarrow$  $\emptyset \sim$ e
      II  $\leftrightarrow$  a
      I  $\leftrightarrow$  o

Masculine is the default gender and is not specified in the lexical entry. Redundancy rules supply class features for the feminine (II) and the $\emptyset \sim e$ class (III). Some -e nouns belong to this last category (III) due to syllabification reasons (e.g., *rifl*e 'rifle'). However —and this is the crucial departure from earlier proposals— Harris suggests a new class IV whose Vocabulary Item is *-e*:

(32)  III  $\leftrightarrow$  $\emptyset$
      IV  $\leftrightarrow$  e

Within this category we find words like *pos*e 'pose', where -*e* cannot be derived from any epenthetic process since voiceless alveolar sibilants are permitted (and in fact very common) in coda position. This new morphological theme position solves the problem of the status of this vowel.

A different question arises with regard to plural formation. In class III words, the contact between the stem consonant and the plural marker yields an illicit sequence, as in *jamón* 'ham' > *\*jamons* 'hams'. To generate the correct plural *jamones*, -*e* must be inserted. Is it epenthetic? As this vowel is also present in contexts with no syllabification problems (*ley* 'law' > *leyes* 'laws', cf. *coméis* 'you eat'), Harris concludes that there is a new context for Vocabulary Insertion in class III plurals:

$$(33) \quad e \longrightarrow \left\{ \begin{array}{l} \text{III} \\ \text{IV\_[plural]} \end{array} \right\}$$

The *morphological* theme position licenses the use of *-e*.[26] This new analysis has a clear advantage, as it solves plural formation. If this position is not assumed, and instead we believe in a purely phonological solution, the syllabification rule applying to words like *grande* 'big' (*\*grand*) would generate an output *\*jamonse*, as the context for epenthesis arises after the plural morph insertion. Otherwise, two allomorphs for plural must be postulated. The 'templatic' solution offers a more straightforward answer for this problem and also unifies different phenomena.

How can we connect this proposal to nasal depalatalization? The idea is that if the underlying form for *desdén* were /desdeɲ/, the rule of vowel epenthesis would apply and would yield a form [desðéɲe]. As this is not the case, it is assumed that Readjustment Rules account for the phonological differences between the verb and noun forms.

To conclude, Harris assumes that for non-canonical endings the word marker feature is simply stored with the lexical entry. That is, a word like *día* 'day' has a specification that indicates its belonging to class II despite its being masculine.

### 3.1.2.2 Bonet (2006)

Bonet (2006) makes a classification of the Spanish nominal system that incorporates the formal machinery of Mascaró (2007) and Bonet et al. (2007) in assuming that phonology, due to the interaction between PRIORITY and RESPECT, controls allomorph selection, as we previously saw in 3.1.1 for Catalan.

Bonet assumes for Spanish a specific markedness hierarchy dependent on gender:

(34)　　Feminine　　{a>e,Ø}
　　　　Masculine　{o>e,Ø}

Thus, inputs are assumed to contain a root plus the multiple gender exponents depicted in (34). Default nouns, like *loro* 'parrot' and *casa* 'house', have the following representation:

(35)　　Feminine　　/kasa + {a > e,Ø}/
　　　　Masculine　/loɾ + {o > e,Ø}

The constraint PRIORITY assigns the default exponent when faced with inputs like those in (35), as we see in (36) (Bonet, 2006, p. 328):

---

[26] "I emphasize that this environment is entirely morphological and that the vocabulary insertion operation takes place in the Morphological Component of the grammar" (Harris, 1999, p. 70).

(36)     *casa* 'house'

| /kas + {a > e,Ø}/ | RESPECT | PRIORITY |
|---|---|---|
| a. ☞ kása | | |
| b.    káse | | *! |
| c.    kás | | *! |

PRIORITY thus favors the unmarked allomorph. Nevertheless, some words surface with marked endings because they are specified as diacritics in the root. This is what happens with *-e* words like *clase* 'class' (Bonet, 2006, p. 328):

(37)     *clase* 'class'

| /klas$_e$ + {a > e,Ø}/ | RESPECT | PRIORITY |
|---|---|---|
| a.    klása | *! | |
| b. ☞ klase | | * |
| c.    klás | *! | * |

Even though it violates PRIORITY, *clase* has an *-e* ending because RESPECT favors candidate b. over the others, since it maintains the lexical specification of the root in the output.

Bonet (2006)'s analysis is important due to the morphological structure posited for words that had been previously analyzed as undergoing epenthesis. Harris distinguished between the vowel *-e* in *clase*, lexically specified, and the vowel *-e* in *madre* 'mother', a product of epenthesis. Bonet, instead, conflates the two kinds of words into the same group, as she argues convincingly —through the analysis of synchronic epenthetic processes— that final epenthesis is no longer active in Spanish anymore (although it was so at previous stages in the evolution of Spanish). The assumption, thus, argues for the same input in *clase* and *madre*, i.e., /klas$_e$ + {a > e,Ø}/ and /madr$_e$ + {a > e,Ø}/.

As the vowel -*e* is always morphological in nature, it can also be used in plurals without resorting to epenthesis. Therefore, Bonet assumes that in a word like *panes* 'breads' -*e* surfaces due to the subcategorization imposed on *pan* 'bread' (input: /pan$_{\varnothing,pl:e}$ + {o > e,$\varnothing$}/). *Fans* 'fans', instead, does not have any diacritic in the root that imposes -*e* in the plural, and thus we find the form *fans* 'fans' (input: /fan$_\varnothing$ + {o > e,$\varnothing$}/).

Bonet (2006)'s reinterpretation of epenthesis will prove to be crucial for analyzing the nominal system of Pallarese Catalan, as I also assume that epenthesis applies only internally and word-initially (see 4.1.2 for more details). In my proposal, though, no special diacritics will be added to the root, as inflectional endings are fully stored with lexical entries as part of the stem in non-regular cases.

### 3.1.2.3   Bermúdez-Otero (2006b, 2007, 2013)

The basic innovation provided by Bermúdez-Otero (2006b, 2007, 2013) lies in the storage of fully inflected stems as lexical entries (i.e., with their corresponding theme vowels).[27] That is, while most generative analyses of Spanish, following Harris, argue for a model of morphology that only stores roots in the lexicon, Bermúdez-Otero considers that theme vowels are also stored as part of lexical entries. [28]

We have just seen that Bonet (2006) posited subcategorization requirements for certain roots, such as /klas$_e$/ 'class'. While this has an effect in the final output, forcing /klas$_e$/ to surface as [kláse] (and not as *[klás], syllabically well-formed), the theoretical machinery differs substantially from that of Bermúdez-Otero, where the lexicon stores not roots but stems and, therefore, the lexical entry for the word *clase* is /klas-e/.

The full nominal system as divided up by Bermúdez-Otero is the following:

---

[27]The term 'theme vowel' is used by Bermúdez-Otero to refer to inflectional endings in the nominal system.

[28]Some authors like Pena (1999) and Pensado (1999) have argued as well that theme vowels are stored in lexical entries and are subsequently deleted in derivatives. Similar proposals have been made for other Romance languages like Italian (see Scalise 1983).

(38)

|   |   |   |   |   |
|---|---|---|---|---|
| 1) | *o*-stems (masculine) | /liβɾ-o/ | *libro* | 'book' |
| 2) | *a*-stems (feminine) | /kas-a/ | *casa* | 'house' |
| 3) | *e*-stems (both genders) | /liβɾ-{e,Ø}/ | *libre* | 'free' |
|   |   | /kɾuθ-{e,Ø}/ | *cruz* | 'cross' |
|   |   | /klas-e/ | *clase* | 'class' |
| 4) | athematic stems (both genders) | /xeɾsej/ | *jersey* | 'pullover' |
|   |   | /biɾus/ | *virus* | 'virus' |

All nouns are thus stored with their corresponding theme vowels, except for the athematic class. For masculine and feminine, *o*- and *a*-stems are the default, respectively, although *e*-stems are also quite common for both genders. Within this group, Bermúdez-Otero makes a distinction between ordinary *e*-stems (*cruz* and *libre*) and *e*-only stems (*clase*). As can be seen in (38), morphology provides two allomorphs in the input for ordinary *e*-stems. Phonology decides between them, opting for -*e* when the final coda (or coda cluster) of the root is phonotactically illicit, as in *libre*, and for -*Ø* when it can be syllabified, as in *cruz*. Crucially, however, this vowel is in no way linked to epenthesis: it is part of the inflectional input and, as such, morphological (as in Bonet 2006; this theme vowel is also used in the plural, as in *cruces* 'crosses'). In *e*-only stems like *clase*, by contrast, there is no allmorphy and the vowel is always present in the lexical entry.

Stem-storage is not a trivial assumption and has many implications for the nominal system. From an empirical point of view, it is important to look at the behavior of derivatives like *manaza* 'big hand' in (39):

(39)   *Base word*       *Augmentative*
      /man-o/ 'hand'    /man-aθ-a/ 'big hand'

In a root-based system, suffixes like -*aθ*- are directly attached to roots, and thus it is clear why the inflectional ending -*o* does not appear in the derivative. In a stem-based system, by contrast, we need to explain the non-appearance of the inflectional ending in the output, i.e., *[man**o**aθa], and Bermúdez-Otero advocates a morphophonological stem-final vowel deletion rule before derivative suffixes.

Using Stratal Optimality Theory (Kiparsky 1998; Bermúdez-Otero 1999), a cyclic version of Optimality Theory, Bermúdez-Otero argues that the absence or presence of 'stem-final' vowels in derivatives is related to the morphological structure of different noun classes and the domain of attachment of certain suffixes. Some suffixes are stem-level while others are word-level. This evidence is provided by diphthongization, as -*dad* is attached to the stem but -*ísim-o* to the full inflected word

(diphthongization is a stem-level phonological process). In both cases the stem-final vowel deletion rule applies, as shown below:

(40)   a)   *b**on**-dad*      'goodness'

   b)   *b**ue**n-ísim-o*   'best'

If *-ísim-o* were a stem-level suffix, the correct output should be *\*bonísimo*. As this is not the case, a distinction between suffixes behaving like *-dad* and others behaving like *-ísim-o* must be postulated.

   The problem for Bermúdez-Otero's analysis is that he is obliged to stipulate that stem-final vowel deletion is noniterative, as shown by nouns ending with a sequence of unstressed vowel + theme vowel (adapted from Bermúdez-Otero (2013, p. 44), as shown below:

(41)

| | Base | | Derivative | |
|---|---|---|---|---|
| | [éɾo-e] | 'hero' | [eɾo-ísm-o] | 'heroism' |
| | [kɾáne-o] | 'craneum' | [kɾane-áθ-o] | 'craneum.AUGM'/'headbutt' |

If stem-vowel deletion were iterative, we would find the unattested forms *\*[eɾ-ísm-o]* and *\*[kɾan-áθ-o]*. The fact that even in adverbial derivatives like *adelantar* 'overtake' (base: *delante* 'ahead, \*adelant**ea**r) the stem-final vowel deletion applies leads Bermúdez-Otero to conclude that this rule is blind to the morphological affiliation of its target.[29] Therefore, it should apply to bases like *héroe* and *cráneo* once the theme vowel has been already deleted, but the fact that stem-final vowel deletion is noniterative explains its underapplication in these contexts.

   Besides this, there is an open issue which has never been satisfactorily resolved in Spanish noun classes, the status of word-final *-s* in *e*-stems and athematic nouns. Bermúdez-Otero distinguishes three different groups (adapted from Bermúdez-Otero 2006b, p. 298-299):

(42)

| | SG | PL | DIM | DIM-PL | |
|---|---|---|---|---|---|
| a. | iris-∅ | iris-e-s | iris-it-o | iris-it-o-s | 'iris' |
| b. | virus | virus | virus-it-o | virus-it-o-s | 'virus' |
| c. | crisi-s | crisi-s | cris-ecit-a | cirs-ecit-a-s | crisis |

*Iris*-like words are *e*-stems and follow the morphological requirement that selects *-e* in plurals of this class (as in Harris 1999). In the case of *virus*, the distinction between this form and derivatives like *viral* 'viral' is due to the root-based nature of

---

[29]Unlike Harris, he does not consider thematic the vowel endings of adverbs and prepositions.

the latter (a bound root allomorph). It is important to note that there are therefore two roots, *virus-* and *vir-*, and *-us* is not an independent inflectional marker.[30]

Finally, nouns like *crisis* are considered pseudoplurals (note the difference between *virusito* and *crisecita*). They lose their pseudoplural ending *-s* when confronted with suffixation, as in *crisecita*. An important observation regards another category of pseudoplurals like *Carl-o-s* which, according to the author, belong to the regular *-o* class as shown by the diminutive *Carl-**it**-o-s*. The claim is that in this regular category *-it* is not a suffix but an infix (*-it* infixation is found in other contexts, e.g., [aθukítaɾ] from [aθúkaɾ] 'sugar'), whereas the first group takes the suffix (see Bermúdez-Otero 2007 for a thorough analysis of Spanish pseudoplurals).

The analysis of athematic nouns (in particular this last distinction between classes of pseudoplurals) and the morphological structure outlined for regular nouns and adjectives will be important for the proposal adopted for Pallarese Catalan presented in 7.1, as my analysis assumes stem-storage for all non-default elements (but, crucially, not for the default ones).

## 3.2 Number

Neither in Catalan nor in Spanish faces the analysis of number major problems. Both languages distinguish between singular and plural and, as part of Western Romania, they pluralize by adding the suffix *-s* after the inflectional ending, as shown in (43):

(43)   *gossa*   'female dog'   >   *gosses*   'female dogs'   (Catalan)
       *vaso*    'glass'        >   *vasos*    'glasses'       (Spanish)

While there is no doubt that the plural suffix is underlyingly /s/ in Spanish (Spanish does not contain voiced fricatives), in Catalan it could be either /z/ or /s/ as we find both forms depending on the phonological context.[31]

In Central Catalan, the word *gosses* is pronounced as [gósəs] in isolation or before a voiceless consonant but as [gósəz] before a vowel or a voiced consonant, as in *gosse*[z] *alegres* 'happy female dogs'. No general conclusions about the underlying nature of the plural morph can be drawn from these data for different reasons, however. On the one hand, final devoicing and voicing assimilation, both general phonological processes in Catalan, generate outputs with a voiceless sibilant in word final posi-

---

[30]See section 7.1 for the same assumption in Pallarese Catalan.

[31]Sibilants in Spanish can also be (partially) voiced in certain contexts, e.g., when they appear as codas in morpheme boundaries and are followed by a vowel (e.g., *los **a**migos* 'the friends'), but there is no phonological contrast between voiced fricatives and their voiceless counterparts, as opposed to Catalan.

tion or before voiceless consonants. On the other hand, sonorization of fricatives, another regular process in Catalan by which voiced and voiceless coda fricatives are neutralized before words starting with a vowel, always yields [z] on the surface (voicing assimilation produces the same effect before voiced consonants). Therefore, several phonological phenomena prevent us from knowing if the plural morph is /z/ or /s/ underlyingly. However, Mascaró (1986) provides evidence in favor of assuming /z/ to be the plural morph by comparing it with the reflexive/impersonal clitic *es*. In a sentence like *Allí **s**'hi arriba a peu* 'We get there on foot', the impersonal clitic, which appears between vowels, is voiceless. In the same phonological context, by contrast, the plural morph remains voiced, as in *Les gosse*[z] *hi van de pressa* 'The female dogs get there quickly'. Thus, we must conclude that it is voiced in the underlying representation, because otherwise it would surface as [s] in intervocalic contexts.

As we have already seen in this chapter, phonology can have an effect on the realization of certain elements in the nominal system. This is what we see in Catalan for masculine nouns ending in a sibilant, like *pes* 'weight' > *pesos* 'weights', which use a gender morph to avoid an OCP-sibilant contact, or Spanish -*es* plurals, such as *jamón* 'ham' > *jamones* 'hams', where -*e* is considered epenthetic under certain analyses. Some of the problems posed by these plurals have already been analyzed and others will be further discussed in 7.1.

Other problems regarding number which have not been dealt with so far are:

- Defective nouns (*singularia* and *pluralia tantum*)

- Invariable nouns and adjectives (e.g., *cactus* 'cactus' > *cactus* 'cacti')

A full exploration of the (purely) morphosyntactic and/or semantic puzzles that this set of nouns pose remains outside the scope of this dissertation. However, we will discuss the consequences for exponence of these mismatches when we analyze the nominal system of Pallarese Catalan in 7.1.

# PHONOLOGICAL SPELL-OUT: AN AUTOSEGMENTAL PROPOSAL

## 4.1 Phonology after morphology in noun inflection

Inflectional systems of the world's languages use different strategies to create phonological distinctions for certain morphosyntactic values. Romance languages like Catalan or Spanish, as we have already seen in chapter 3, make use of morphological markers at the right edge of the word to distinguish masculine from feminine nouns, for example. Other languages, by contrast, do not use segmental material and, instead, resort to suprasegmental information for the same purposes. One of the most prominent and well-studied features belonging to this category is tone, implemented by some languages. An example of the morphological use of tone can be seen in (1), with examples from dialects of Limburg Dutch (adapted from van Oostendorp 2005, p. 108):

(1)  | *neuter* | *feminine* | *masculine* | |
     |----------|------------|-------------|---|
     | wíís | wîs | wîzə | 'wise' |
     | dóúf | dóùf | dóùvə | 'deaf' |
     | láám | láàm | láàmə | 'lame' |

Adjectives in Limburg Dutch make a distinction between neuter and feminine forms only by tone. As shown in (1), the neuter bears a level high tone whereas the feminine has a falling tone.[1] All the segmental material being equal, only tone is used to

---

[1]The terms "level high tone" and "falling tone" are employed by van Oostendorp.

distinguish syntactic gender. Therefore, we can conclude that tone plays the same role as inflectional class markers. Even though they serve the same purpose, tone is special in one respect: it is a semi-dependent feature. That is, while inflectional vowel endings in Romance are independent segments which can be attached to a root, tone needs to be linked to a given segment (or group of segments) to be interpreted phonetically.

Autosegmental Phonology (Goldsmith, 1976) is a particularly adequate framework for the analysis of tone. The fact that tone and melodic content are dissociated (as they are in different tiers linked through the skeleton) easily explains certain tonal phenomena like spreading or multiple association of tones. But Autosegmental Phonology has been widely used also in the analysis of other phonological phenomena such as vowel or nasal harmony. Segments consist of bundles of features and, therefore, specific features are independent from the timing unit they need to be attached to to be pronounced. In fact, if a given feature is not linked to the skeleton it remains floating and it does not receive phonetic interpretation. This well-accepted assumption in phonological theory will prove to be crucial to the understanding of all the singularities of Catalan nominal inflection, as we will see throughout this chapter. The idea pursued in this thesis is simple: nominal inflection in Catalan must be understood as a process of featural affixation, following the insights in Akinlabi (1996). That is, Catalan gender exponents like -*a*, default for feminine, are not properly segmental morphemes; instead, masculine and feminine gender values are associated with floating features. How these features end up associated with a specific position depends on the morphological structure that we assume for the nominal system, which is detailed in 4.1.1.

Akinlabi presents a featural affixation analysis based on Chaha, a Gurage language of Ethiopia. Chaha's inflectional system is, predictably, very different from that of Catalan, but there are nevertheless some similarities which may help the understanding of the theoretical proposal made in this thesis. In Chaha, the 3rd singular masculine object is expressed by labializaion, which surfaces on the 'rightmost labializable consonant' of the stem. However, only labial and dorsal consonants are labializable, not coronal, as (2) shows (adapted from McCarthy (1983, p. 179):

(2)

  a. **Rightmost consonant of the stem is labializable**

     *Without object*   *With 3rd m. sg. object*

     näkäb     näkäb$^w$     'find'

  b. **Medial consonant of the stem is labializable, final is not**

     näkäs     näk$^w$äs     'bite'

  c. **Only the leftmost consonant of the stem is labializable**

     mäsär     m$^w$äsär     'seem'

  d. **No labializable consonant**

     sädäd     sädäd     'chase'

The examples in (2) show the phonological conditioning imposed on Chaha's verbal morphology. When all the consonants of the verb are labializable, as in (2a), the 3rd singular masculine object is realized at the rightmost consonant. When the last consonant is not dorsal or labial, as in (2b), the [round] feature skips it and is linked instead to the next rightmost consonant; if neither of those consonants is labializable —(2c)— but the leftmost consonant is, then the leftmost consonant is labialized. Finally, when the root only contains coronal consonants, the [round] feature associated with the 3rd singular masculine object does not surface.

Akinlabi argues, in his OT analysis, that a morphological alignment constraint ALIGN (3M. SG., R, STEM, R) ("the right edge of the 3m. sg. must be aligned with the right edge of the stem") is overranked by a feature co-occurrence constraint *COR/LAB ("if [coronal], then not [round]"), which explains why the [round] feature surfaces in different positions in the stem or it does not surface at all. I will not go into the details of the proposal here; the relevant point is that the parsing of underlying input features associated with morphosyntactic values is constrained by phonological requirements, which may license the linking of those features. As Akinlabi points out, "the surface realization of a featural affix is crucially dependent on possible licensors" (p. 251). This supports the late insertion idea defended in 1.1 and shows that, even if morphology connects morphosyntactic features to phonological features, it is in the phonological module where the surface realization of some morphemes is finally determined.

Chaha data cannot be compared to Catalan for obvious reasons: whereas in Chaha features are spelled out in other segments (the [round] feature needs to be attached to the rightmost consonant that allows its anchorage), in Catalan we always find a gender exponent at the right edge of the word and, thus, it seems more straightforward to treat nominal inflectional markers as full segments. If we assume,

instead, that the spell-out of gender consists of floating features (as I am doing), how do we prevent these features from being linked to the already-existing skeletal positions of the root, thus changing their featural inventory? The general idea is that an independent morphological position attracts these features and licenses their linking to this position. I will label this autonomous morphological position 'theme position' following Oltra-Massuet (1999)'s analysis of Catalan morphology, that was previously shown in 3.1.1. In the following section, 4.1.1, I will review the role of the theme vowel in the nominal system of Catalan and the implications it has for the autosegmental analysis defended in this dissertation.[2]

## 4.1.1   Theme vowels and gender spell-out

Romance languages like Catalan use arbitrary vowel markers known as 'theme vowels' which are "empty constituent<s> that appear next to a root in order to build a stem or a base to which inflectional affixes can be attached to build a well-formed word" (Oltra-Massuet 1999, 19). Traditionally, the notion of theme vowel has been connected to the verbal morphology, but in the present analysis, as in Oltra-Massuet's, it is extended to other morphosyntactic categories. Theme vowels are not semantically, syntactically or phonologically motivated, but they are an essential component of the structure of verbs, *nouns* and *adjectives*, as they indicate the class into which each category falls.

What is the status of theme vowels in the overall structure of the nominal system, then? First of all, to understand the role that the theme vowel plays as part of the stem of a word we must determine how words are created. In this thesis, a theoretical model is assumed in which roots are acategorial and need to be attached to a category defining head, *n* in this case (Marantz 1997, 2001). As Alexiadou (2001) points out,

> (...) what we think of as the syntactic categories N, V and A are actually morphological categories created by the syntax, i.e., post-syntactically realized. Certain functional heads play a central role in defining a domain for syntax and for phonological and semantic interpretation. It is syntax that creates the words, and it is syntax that makes sense of the constraints on word formation.

It follows from this assumption that the functional head *n* creates a lexical category when attached to a root. Although this head is, most of the time, phonologically

---

[2]See chapter 5 for the formal OT analysis of Pallarese Catalan gender exponence.

null, it can be realized by some morphemes, especially in derivational contexts (suffixes like *-ció*, as in *concepció* 'conception' or *-ment* as in *casament* 'marriage', for example).

We have already seen in 3.1.1 that Oltra-Massuet (1999) adopts the non-lexicalist approach to building words described above. In addition, she proposes a further step for Catalan inflection that consists of a well-formedness requirement imposed on every syntactic functional head, like little $n$ or little $v$. This is exemplified below for nouns:

(3)     *Adjunction of categorizing little n*

```
        n
      /  \
   root   n
```
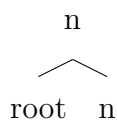
(4)     *Well-formedness condition: every functional head requires a theme position*

```
        n
      /  \
   root   n
        /  \
       n    Th
```
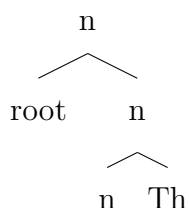
This theme position actually captures the morphological essence of nominal class features in Romance, which do not participate in any syntactic phenomenon (Alexiadou, 2004) but are nevertheless the phonological realization of grammatical gender. It is obvious that a position of some kind must be assumed for Romance languages because the expression of gender falls in a vocalic element at the rightmost edge of the word. We can label it theme vowel, as in (4), or not, but it is impossible to explain Romance inflectional systems without resorting to an autonomous morphological position of some kind. The advantage of Oltra-Massuet's proposal is that it develops a formal structure for this position.

The well-formedness condition that we see in (4) is, as I said, imposed on every syntactic functional head. In verbs, the core of Oltra-Massuet's analysis, this gives rise to a very innovative proposal. According to the author, apart from little $v$, Tense, Aspect and Mood would project theme positions. This can be seen in the following example (adapted from Oltra-Massuet 1999, p. 30):

(5)     a profund **i**   r     **i**  Ø     **a**  z
        $v$ $\sqrt{root}$   Th Mood Th Tense Th Agr
        *aprofundiries* 'you would deepen'

Bold letters correspond to little $v$, Mood and Tense theme vowels, following the well-formedness requirement of (4). For nouns, things are a bit different, as these functional projections are not part of the nominal domain, except for deverbal constructions (in this case, the verbal root will become a noun due to the merge of little $n$ and the lack of c-commandment of Tense/Aspect/Mood by little $v$, which would yield a verb). There should also be more than one theme vowel in derivatives, but when suffixes are joined to nominal stems the theme vowel of the stem does not surface, as can be seen in (6):

(6)  *cas**a***  'house'  >  *caseta*  'little house'  (\**cas**a**eta*)

If, as argued by Oltra-Massuet, every functional projection requires a theme position, how is it that only one theme vowel is found in derivatives like *cas-et-**a***? Several options are available depending on the theoretical assumptions one makes. Bermúdez-Otero (2013) reviews these options departing from the denominal adjective *nuboso* 'cloudy' (from *nub**e*** 'cloud'). As he points out (Bermúdez-Otero, 2013, p.22), the first possibility is to assume that the root is directly attached to an adjectival head $a$, as in (7):[3]

(7)



It is clear, then, that only one theme vowel is realized because little $a$ is merged with a bare root and, thus, there is only one theme position. If we do not attach little $a$ to the noun *nube*, which itself contains a theme position, there will only be the theme position provided by the adjectival head.

Things get a little more complicated if we assume the derivation to be noun-based. In that case, an outer $a$ head is attached to an inner $n$ head (i.e., the stem *nube*). Three possible scenarios regarding the theme position in this derivational process need to be considered (following Bermúdez-Otero 2013).

One possibility is advanced in Oltra-Massuet & Arregi (2005), where both $n$ and $a$ heads have a theme position, but the inner position is not realized, as in (8):[4]

---

[3]The morphological structures shown in (7), (8), (9) and (10) are all extracted from Bermúdez-Otero (2013, p.22-23).

[4]Oltra-Massuet & Arregi (2005) follow Oltra-Massuet (1999) in the assumption that there is a requirement to generate a theme position for every syntactic functional head at the morphological component.

(8)

```
                        a
              ┌─────────┴─────────┐
              n                   a
          ┌───┴───┐           ┌───┴───┐
        √NUB      n       [a,-os-]  [Th, -o-]
              ┌───┴───┐
          [n, -Ø-]  [Th, -Ø-]
```
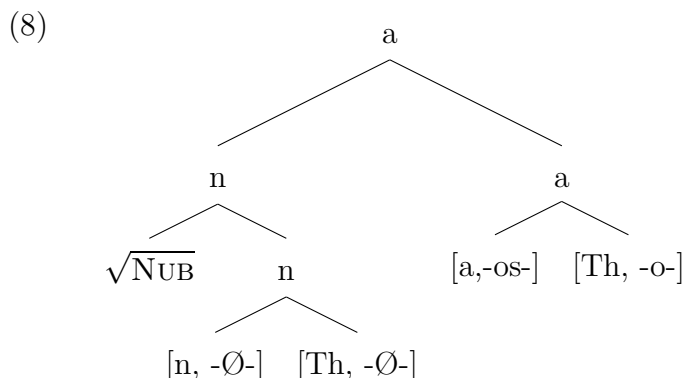
The authors claim that "the theme position of *n* or *a* is not realized overtly if this head is not the highest one in the structure. For instance, when the head *a* is realized as *-ist*, its theme is *-a* only if this head is the highest one (e.g., *novel-**ist-a** * 'novelist'), but not otherwise (e.g., *novel-**íst-Ø**-ic-o* 'novelistic'). On the other hand, the theme adjoined to *v* is always realized phonologically (e.g., *comunic-**a**-ción* 'communication'), with idiosyncratic exceptions." (Oltra-Massuet & Arregi 2005, footnote 8).[5]

A second possibility would be the adjunction of a theme position only in the outer head, but this would go against the well-formedness requirement of (4):

(9)

```
                    a
            ┌───────┴───────┐
            n               a
        ┌───┴───┐       ┌───┴───┐
      √NUB   [n, Ø]  [a,-os-] [Th, -o-]
```
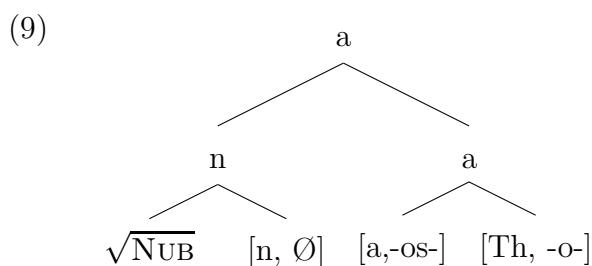
There is one last option, the one pursued by Bermúdez-Otero:

(10)

```
                        a
              ┌─────────┴─────────┐
              n                   a
          ┌───┴───┐           ┌───┴───┐
        √NUB      n       [a,-os-]  [Th, -o-]
              ┌───┴───┐
          [n, -Ø-]  [Th, -**e**-]
```
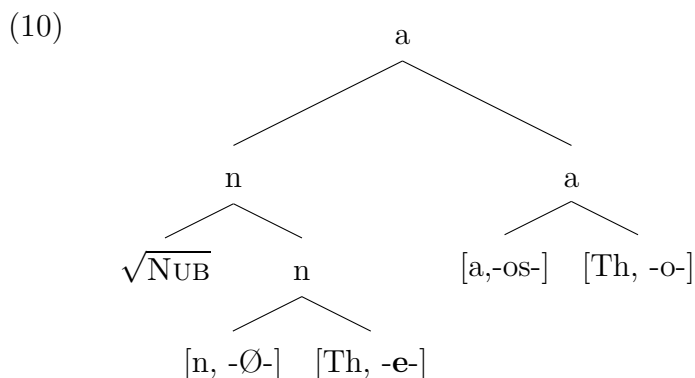
In this latter case the theme positions of both heads are overtly realized. The

---

[5]See Bermúdez-Otero (2013, footnote 19) for a criticism on this view due to conceptual incoherences with the tenets of Distributed Morphology.

morphological structure of (8) and (10) is exactly the same, but their phonological realization is not. Bermúdez-Otero argues that there is a phonological rule deleting the stem-final vowel of the inner head (along the lines of Scalise 1983; see 3.1.2), yielding the following derivation for the adjective:

(11)    [[nub-**e**]os-o]   >   [nubóso]

Of the four possible structures sketched above, those belonging to Oltra-Massuet & Arregi (2005) and Bermudez-Otero's analysis seem the most plausible. A strictly root-based approach is not adequate because it does not capture the syntactic category the adjective is derived from (that is, in Spanish the adjective *nuboso* is based on the noun *nube*. If roots are acategorial, the adjunction of little *a* to the root misses this point). On the other hand, a stem-based approach like (9) has to stipulate that only the outer head gets a theme position. If nouns have theme vowels and the adjective is derived from a noun, there is no apparent reason why this position should not exist.

Oltra-Massuet & Arregi (2005) and Bermudez-Otero's proposals ((8) and (10), respectively), although more empirically adequate, face some problems too. If the theme position is not realized, we have to explain why. Why is there only a theme vowel in the outer domain in (8)? The structural reasons for this are not clear, and the assumption only describes the facts of the language. If, by contrast, a theme vowel is present underlyingly, we have to explain the opposite, i.e., its non-appearance in the output and the mechanism that deletes the stem-final vowel. Bermúdez-Otero's stem-final vowel deletion rule explains why the adjective *nuboso* in (11) does not contain the theme vowel of the inner head but it has to stipulate that this rule is noniterative. For the sake of clarity, I repeat in (12) what was presented in 3.1.2 to exemplify this:

|  | *Base* |  | *Derivative* |  |
|---|---|---|---|---|
| (12) | [éɾo-e] | 'hero' | [eɾo-ísm-o] | 'heroism' |
|  | [kɾáne-o] | 'craneum' | [kɾane-áθ-o] | 'craneum.AUGM'/'headbutt' |

The stem-final vowel deletion rule does not distinguish the morphological status of the vowel it has to delete for the reasons set forth in 3.1.2, and thus the vowel of the root should not surface either. Therefore, instead of the derivatives *her**oi**smo* and *cran**ea**zo* we should find the forms *\*herismo* and *\*cranazo*, with no vocalic contact. This leads Bermúdez-Otero to conclude that the rule must be noniterative, because otherwise it would overapply.

Both proposals need to stipulate at some point, and it does not seem easy *a*

*priori* to favor one option over the other. Under the present analysis, though, only the vowel deletion rule can account for derivational processes in the nominal system, as we will see in chapter 7 that non-regular nouns and adjectives are lexicalized. This means that their theme vowel is already spelled out in the lexical entry, i.e., its phonological content is fully specified. When a suffix is added to create a diminutive, for example, the vowel is already realized but it does not appear in the output (*llor-o* 'parrot' > *llor-et* 'little parrot', \**llor-o-et*), which implies that there must be a rule that prevents this inflectional marker to surface in derivational contexts in that position, i.e., a rule that deletes it, because it was already spelled out.

In any case, what really matters in my analysis and, thus, needs to be clearly stated is the spell-out process of gender and how it is connected to the morphological structure just posited. In fact, until now, we have analyzed the morphological structure of the nominal system and the theme position that serves as location for inflection, but no reference to the exponence of this position has been made. Once the morphological theme position has been clearly defined, which is crucial to understand the phonological spell-out process of gender, we can proceed to the formal analysis of Pallarese Catalan nominal inflection.

As I have already mentioned, I am assuming late insertion throughout this dissertation; that is, exponents are only accessible after morphosyntactic operations. This implies that the input to phonology is the output of morphology, a crucial step in the analysis presented here. In Bye & Svenonius (2012) words, we can define morphology as follows:

> Somewhere between the system of syntax-semantics and the system of phonology, there is an interface in which *representations legible to the one system are mapped to representations legible to the other*. (Bye & Svenonius, 2012, emphasis mine)

Morphology is thus a 'transfer system', an idea developed in Bermúdez-Otero (2012) as well. Under Bermúdez-Otero's view, morphology cannot manipulate syntactic nor phonological features, it can only give instructions for insertion. Regarding gender, morphology's duty is thus to associate a [±fem] value with phonological features. We have to determine, however, *how* these features are spelled out, and what is assumed in this dissertation is that only floating phonological features are associated with gender. That is, they have been translated into material that can be interpreted in the phonological component, but they are not linked to any timing unit (a mandatory step before these features can be interpreted phonetically). Therefore, regarding inflection, the duty of phonology is to insert association lines

between phonological features and skeletal positions. If features are not linked via association lines, they remain floating and, as a consequence, unpronounced.[6]

Once we have delimited the tasks phonology and morphology accomplish, we can proceed to the formal implementation of the proposal. Keeping in mind the morphological structure posited in (4), Vocabulary Items are inserted for both gender features and the theme position in Pallarese Catalan:

(13)    a.    Underspecified vowel in the theme position:
            Th ↔ V (*à la* Roca 2010)

        b.    Association of gender values with phonological features:

            [+fem]   ↔   [+low]           (= /a/)
            [–fem]   ↔   [+labial, –high]   (= /o/)

As can be seen in (13), Pallarese nominal inflection consists of a two-step process. On the one hand, the theme position is filled with an underspecified vowel, (13-a). This means that a [–cons] feature is associated with a timing unit but the quality of the vowel is not specified. This is what Bye & Svenonius (2012, p. 442) label a 'featurally deficient segmental root node', and it is considered one of the options for deficient lexical entries which do not constitute the default morphemic exponence, i.e., a segment or a string of segments.[7] Roca (2010) proposes the same underspecified V-slot for the Spanish verbal system, but I extend it here to all theme positions because it captures the idea that there is a fixed morphological space for inflection, no matter what phonological (vocalic) feature may end up there.[8]

On the other hand, as shown in (13-b), morphology provides associations between gender values and phonological features, but these features are not linked to any timing unit, i.e., they are floating.[9] These features will never be pronounced if they are not associated with an X-slot, but the morphological theme position already provides one and attracts them. Adding an association line is less costly than

---

[6]The idea of analyzing Catalan gender inflection in autosegmental terms was suggested to me by Ricardo Bermúdez-Otero, whom I would like to gratefully thank, as much of the analysis presented here is inspired on his suggestions.

[7]Other options include moras, syllables, feet, etc. Bye & Svenonius (2012) use the term 'deficient' in this type of lexical entries because they cannot be interpreted phonetically *per se*, i.e., they need 'full segmental content' (p. 442).

[8]Roca (2010)'s analysis, however, introduces a complex set of morphological constraints that provides the correct featural content for each conjugation class theme vowel. As we will see in 5, morphological constraints are not needed in the present analysis of gender inflection, as the featural content provided by Vocabulary Items and phonological constraints can deal with the patterning of the nominal system.

[9]Only the distinctive phonological features for the vowels *-a* and *-o* are provided in (13). The masculine default exponent for Pallarese is -[o], whereas in Central Catalan it would be -[u] due to a different process of vowel reduction.

creating a new structural position (which needs an association line anyway) and is thus preferred. Moreover, the features will not end up in the skeletal positions provided by the root due to feature-identity constraints that advocate for faithful input-output mappings.

One could argue that the fact that the theme position and gender are spelled out in 'opposite directions' is an *ad hoc* explanation. That is, for the theme position, an X-slot is provided with no specifications for place features, whereas [±fem] only gets floating place features as exponents, with no skeletal position to which these can be attached to. The difference in these Vocabulary Items is nevertheless empirically adequate, as it maintains the autonomous morphological space of the theme vowel (it signals the location for inflection), which is given a privileged status in the system, no matter what vowel quality this may have. As for gender, we can treat it as an 'aerial' feature of some kind. Default featural exponents for masculine and feminine are pervasive in the nominal system and used productively in the creation of new forms.[10] They do not need, however, to be stored as full segments due to the independent status of the theme position, which constitutes an optimal anchoring point for them.

All in all, and given the assumptions in (13), the spell-out process for a feminine word like *casa* 'house' looks as follows:

(14)
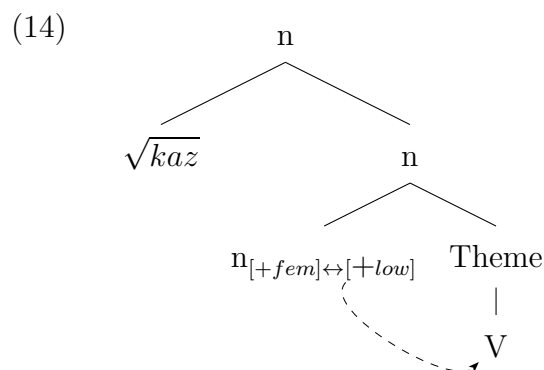


We can see in (14) how the features associated with [+fem] are transferred to the theme position (it is assumed in the present work that gender is stored in little *n*, along the lines of Kramer 2009, 2014; see 4.2.2 for further discussion on this topic, which is far from clear in the literature). That is, morphology has provided the association of [+low] to the [+fem] value of little *n*, but this feature needs to be linked to a timing unit to undergo phonetic interpretation. At this point, the V-slot

---

[10]In the masculine, the morph *-o* is marked with respect to *-Ø*. As we will see in chapter 5, this is a consequence of the constraint ranking because *-Ø* is, in fact, the non-linking of the features associated with *-o*, and, thus, *-o* is still the default marker. Given the distribution of masculine nouns in Catalan it can be misleading to call it default, but it is the default marker in the present analysis.

attracts [+low] as it is already a root node. After this operation, *casa* will be fully interpreted as in (15):

(15)

```
                    n
              ╱           ╲
          √kaz              n
                        ╱       ╲
                  n[+fem]      Theme
                      │           │
                      Ø           a
```
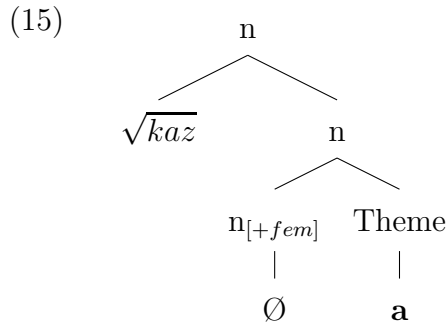
This is the basic machinery that we need to account for the distribution of nominal inflection in Pallarese Catalan. I will argue that the morphophonological analysis assumed here can deal with all the p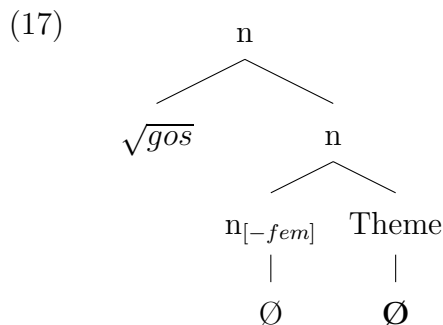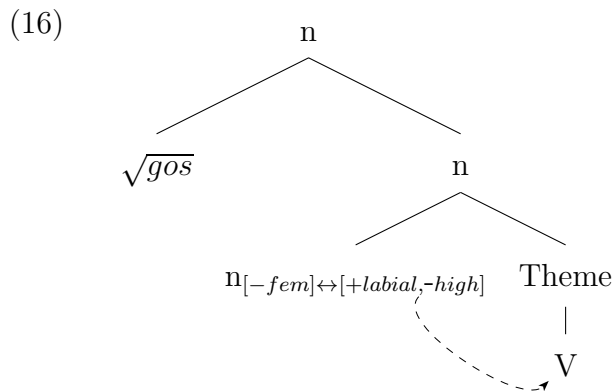honetic variants of Pallarese 3rd person accusative clitics and definite articles, the focus of this dissertation. The details of the OT analysis are presented in chapter 5 for this set of clitics, and are extended to other pronominal cltitics and all classes of nouns and adjectives in chapters 6 and 7, respectively.

Before we move on, there are a couple of things that need to be clarified. First, it is important to understand that the featural exponents shown in (13) can remain floating under certain conditions. This is what happens in the masculine for regular words like *gos* 'dog', as shown in (16) and (17) (the final output):

(16)

```
                       n
                 ╱              ╲
             √gos                  n
                              ╱        ╲
                   n[−fem]↔[+labial,−high]   Theme
                                    ⌐           │
                                     ⌐          V
                                      ⌐        ↗
                                       ⌐ ─ ─ ─
```

(17)

```
                    n
              ╱           ╲
          √gos              n
                        ╱       ╲
                  n[−fem]      Theme
                      │           │
                      Ø           Ø
```

The spell-out process that we see in (16) is the same as in (14), i.e., the features that realize the gender value try to get associated with the V-slot of the theme position. Whereas in the feminine the end result is the pronunciation of -[a], in the masculine we find instead no exponent. This is due to the constraint ranking that will be presented in chapter 5, which favors the realization of the feminine in most contexts but maximally restricts the appearance of -*o*, only licensed when constrained by syllable well-formedness requirements. This is why we do not find it in *gos* but we find it instead in the plural *gossos*, because an OCP-sibilant violation bans the form \**goss* and the use of the featural exponent is preferred over the use of the default epenthetic vowel of Pallarese, -*e* (see the details of the proposal in section 5.1.1).

Second, it is clear that not all words in Catalan have default endings. The idea that I develop in this thesis is that all non-canonical endings are lexicalized. For example, for the feminine word *base* 'base', whose inflectional ending is -*e* and not the default -*a*, we have to consider this noun to be stored as a complex structure with the vowel -*e* realizing the theme position, as in (18):

(18)

$$
\begin{array}{c}
\text{n} \\
\diagdown \\
\sqrt{baz} \qquad \text{n} \\
\diagdown \\
\text{n}_{[+fem]} \quad \text{Theme} \\
| \qquad\quad | \\
\emptyset \qquad\quad \text{e}
\end{array}
$$
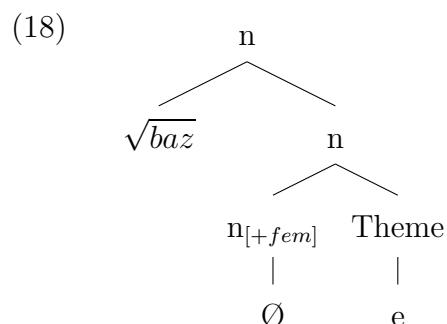
In (18), the regular gender spell-out process will try to apply as in (14), but the fact that the theme position is already occupied by a fully realized segment blocks the linking of [+low] (again, because featural identity constraints disprefer unfaithful input-ouptut mappings). I will not elaborate further on this assumption here because the whole noun system is analyzed in detail in chapter 7, but it is important to keep in mind that this proposal generates only the default endings of the nominal system. All other non-regular elements are lexicalized in the theme position.

Finally, it follows from this proposal that epenthesis cannot be a right-edge process in the nominal system, as it interferes with the theme position (the V-slot always occupies inflectional space). That is, when there is a phonotactic problem, as in the example of *gossos*, mentioned above, the morphological structure favors the use of an inflectional morph instead of a purely epenthetic element. Therefore, vowels that have been usually considered epenthetic, like that of the masculine word

*centr**e*** 'center' ( *\*centr* ), are reanalyzed as being lexicalized. In 4.1.2 these and other cases involving epenthesis in Catalan are reviewed.

## 4.1.2 Epenthesis & Morphological Epenthesis

Several studies have treated Catalan vowel epenthesis from a generative perspective (especially Mascaró 1976; Wheeler 1979, 2005). A complete description of the phenomenon and its domains of application is found in Lloret (2002). All authors agree that epenthesis in Catalan is usually peripheral to the word, although in certain environments it can be intermorphemic. They establish three patterns of epenthesis according to its position: word-initial, word-final and word-internal.[11] My claim, though, is that epenthesis applies only word-initially and word-internally, as I argue that word-final epenthesis interferes with the theme position of both the nominal and the verbal system and, as a consequence, it is avoided.[12]

As I said, in word-initial position, epenthesis is fully productive. Loanwords and native words alike are clear indicators of this process, shown in (19):[1314]

(19)

| | | | |
|---|---|---|---|
| Spielberg | > | [ə]Spielberg | 'Spielberg (film director)' |
| MNAC | > | [ə]MNAC | Catalan art museum located in Barcelona |
| SCHOLA | > | [əskɔ̀la] | 'school' |

In Catalan, *sC-* clusters are not allowed word-initially and therefore an epenthetic vowel is inserted before *-s.* This is the case of *escola* and many other words derived from Latin such as *escriure* 'to write', *esperar* 'to wait', ... The same happened in French or Spanish with these Latin clusters, but not in other Romance languages such as Italian:

---

[11]I will leave aside cases of inter-word epenthesis found in Alguerese Catalan. For a detailed analysis of the phenomenon see Loporcaro (1997) and Lloret & Jiménez (2010).

[12]I reinterpret many cases that have traditionally been analyzed as word-internal epenthesis, especially in the verbal system, as non-epenthetic. See below for the details of the proposal.

[13]Hall (2011) argues that loanword and native-language epenthesis cannot be grouped together. Although it is true that they pattern differently and there are perceptual factors that need to be taken into account in loanword adaptation, the differences among them are not crucial in my analysis. What I want to show is that word-final epenthesis is non-existent either in the native lexicon or in loanwords due to the constraints imposed by morphological structure.

[14]MNAC stands for *Museu Nacional d'Art de Catalunya* 'National Art Museum of Catalonia'. Although it is officially pronounced [mənák], many speakers epenthesize at the left edge of the word. I am using Central Catalan for phonetic transcription (and, therefore, schwa as default epenthetic vowel) in this section because most analyses of epenthesis have been based on this dialect. Moreover, Western Catalan dialects –the macrodialectal group Pallarese belongs to– make use of different vowels depending on their position within the word and therefore the analysis would be more complicated (the vowel *-e* is, though, the default element). See Jiménez (2002) and Lloret & Jiménez (2008) for the different vowel qualities found in the epenthetic proceses of Western Catalan dialects.

(20)

| SCHOLA | > | [skwóla] | (Italian) |
|---|---|---|---|
| | > | [ekɔ̀l] | (French) |
| | > | [eskwéla] | (Spanish) |

Although diachronically this has been an active process from Latin to Catalan, it is not clear whether these words have to be interpreted as being underlyingly /sC-/ or /VsC-/. As pointed out by Wheeler (2005, p. 250), given Richness of the Base (Prince & Smolensky, 2004), both inputs could be inferred by the speakers but the relevant ranking would always yield [əsC-] outputs. The problem is that there is not enough empirical evidence to fully support the claim that the initial vowel in these words is epenthetic, as there are morphophonological alternations between [Ø] and [ə] in derived forms:

(21)

| ***e**sperar* | 'to wait' | [ə] |
|---|---|---|
| *des**e**sperar* | 'to make someone desperate' | [ə] |
| *prosperar* | 'to prosper' | [Ø] |

In Lexical Phonology and Morphology (Kiparsky, 1982b) these differences were accounted for by assigning each prefix to a different stratum. Regardless of the chosen theoretical framework, in forms like *prosperar* it is not clear whether speakers are conscious, synchronically, of the decomposition of the word in the prefix *pro* and the verb *esperar*. That is, they probably do not treat it as compositional but assign a semantic value to the whole structure, not to each component. The verb is stored as /prosperar/, not as /pro + esperar/. In Pons-Moll & Lloret (2014) it is also assumed that the morphological compositionality of these cases is opaque and not productive, as opposed to other prefixed forms shown in (22) (the dialect analyzed there is Majorcan Catalan, which also has schwa as the default epenthetic vowel):

(22)

| *super**e**special* | 'super especial' | [ə] |
|---|---|---|
| *in**e**specífic* | 'unspecific' | [ə] |

What the authors claim is that when the morphological compositionality of the word is 'transparent' (i.e., *inespecífic* clearly means 'not specific' –in+específic–), there is no [Ø]~vowel alternation and epenthesis always applies. They admit, though, that the conflicting data of (21) and (22) do not let us establish a clear input for these sequences. Their claim is, however, that the independent evidence provided by the underapplication of vowel reduction in Majorcan Catalan shows that the speaker unfaithfully maps [Ø] to [ə] in *sC-* words, as detailed below (see also Lloret & Pons-Moll (2016)).

Majorcan Catalan reduces [á], [έ] and [é] to [ə] in unstressed positions, a regular phenomenon in Eastern Catalan varieties:

|      |      | Stressed |        | Unstressed |                        |
|------|------|----------|--------|------------|------------------------|
| (23) | c[á]sa | 'house'  | c[ə]s[ɔ́]ta | 'small house (dim.)' |
|      | caf[έ] | 'coffee' | caf[ə]t[ɔ́]t | 'small coffee (dim.)' |

However, the mid front vowels [é] and [έ] do not reduce to schwa in certain environments, especially in derived forms. Crucially, these vowels must be located in the initial syllable of the stem –(24)– to skip the process, as in other positions –(25)– vowel reduction applies as expected (examples from Pons-Moll & Lloret 2014, p. 7):[15]

|      |      | Stressed |        | Unstressed |                    |
|------|------|----------|--------|------------|--------------------|
| (24) | f[é]sta | 'party'  | f[e]stassa | 'big party (augm.)' |
|      | c[έ]l  | 'sky'    | c[e]let    | 'little sky (dim.)' |

|      |      | Stressed |          | Unstressed |                      |
|------|------|----------|----------|------------|----------------------|
| (25) | cast[é]ll | 'castle' | cast[ə]ll[ɔ́]t | 'small castle (dim.)' |
|      | pap[é]r  | 'paper'  | pap[ə]r[ɔ́]t   | 'small paper (dim.)' |

What is the relation between this phenomenon and word-initial epenthesis, though? The key argument is that, in derived forms of words with initial *sC-*, there is also underapplication of vowel reduction, as shown in (26):

|      |      | Stressed |           | Unstressed |                       |
|------|------|----------|-----------|------------|-----------------------|
| (26) | [ə]st[é]ve  | 'Stephen' | [ə]st[e]vet  | 'small Stephen (dim.)' |
|      | [ə]squ[é]ma | 'schema'  | [ə]squ[e]met | 'small schema (dim.)' |

If these words were analyzed as having an initial vowel in their underlying representation, the output should be *[ə]st[ə]vet and *[ə]squ[ə]met in the derived forms, because the stressed vowel in the base forms would occupy the second position and, therefore, it should be targeted by vowel reduction. As this is not the case and all words within this category are systematically affected by underapplication of vowel reduction, we must assume that /sC-/ is the only possible input for these clusters.

Although we cannot extrapolate this conclusion to all Catalan dialects, Majorcan Catalan vowel reduction patterning seems to support the non-vocalic underlying representation of *sC-* clusters (after all, the phonological systems of Western and

---

[15]See Pons-Moll 2012 for a complete review of Majorcan Catalan vowel reduction.

Eastern Catalan are not that far apart). Even if we cannot determine the nature of this vowel by unambiguous empirical procedures, this process of epenthesis is fully productive and pervasive in these consonantal sequences in loanwords, as recently incorporated words demonstrate:

(27)
> [ə]Skype
>
> [ə]Spotify
>
> [ə]Snapchat

When it comes to word-internal epenthesis, we need to distinguish between the native lexicon and loanwords as well. In native words, non-edge epenthesis is mostly found in the verbal domain in Catalan.[16] This kind of epenthesis is present in three contexts:

(28)   a.   2nd person singular Present Indicative
       b.   Future Indicative
       c.   Conditional

In the 2nd person singular Present Indicative there are three verbs that, in principle, show an epenthetic vowel: *obrir* 'to open' and *omplir* 'to fill' (*IIIb* conjugation) and *córrer* 'to run' (*IIb* conjugation):[17]

(29)   *obr-*[ə]*-s, ompl-*[ə]*-s, corr-*[ə]*-s*

The phonotactic problem originated as a result of the contact between root consonants and the 2nd person morph /z/ (\**obrs*, \**ompls*, \**corrs*) forces the use of a vowel, something that does not happen in other II or III conjugation verbs when there is no syllabification problem, such as *témer* 'to fear' (*tems* 'you fear') or *morir* 'to die' (*mors* 'you die'). Moreover, IIb and IIIb verbs whose root ends in a sibilant also behave like *obrir*, *omplir* and *córrer* because the contact with the 2nd person morph /z/ generates an OCP-sibilant conflict (e.g., *cuses* [kúzəs] 'you sew' or *creixes* 'you grow up' [kréʃəs]). In IIIa verbs the OCP situation is the same when the second person morph gets in contact with the sibilant ending of the stem extension *-eix* (*pateixes* -from *patir-* [patéʃəs] 'you suffer').

---

[16]In the nominal system, internal epenthesis appears in cases that have traditionally been analyzed as already epenthetic in singular forms because of the final consonantal cluster in the root. See below —(35)— for an analysis of these forms.

[17]The classification used for verbal classes is based on Fabra (1956) and Viaplana (1986), where both the 2nd and the 3rd conjugation are split into two subgroups depending on the presence of stem extensions or phonological criteria. I will make use of this terminology throughout the dissertation, also adopted in Perea (2002), although Mascaró (1986) puts IIa and IIb together in one single group.

It is not clear, though, if this vowel —epenthetic at earlier stages of Catalan— has been reinterpreted by actual Catalan speakers as a theme vowel, which is already used for all verbs of the first conjugation (/a/) in the 2nd person singular Present Indicative, as shown below:[18]

(30)    *cremar:*    *cremes*    [krémǝs]    'you burn'

        *parlar:*    *parles*    [párlǝs]    'you talk'

The underlying representation for both forms in (30) must contain a theme vowel (/krem-a + z/ - /parl-a + z/) because otherwise the output of the verb *cremar* would be *\*crems*. That is, the contact between the coda cluster *-rl* and the 2nd person morph in *parlar* would yield an illicit sequence (*\*parls*) which could be repaired by an epenthetic element, but that cannot be maintained for *cremes*, as the sequence *-ms* is admitted in Catalan and the use of a vowel between the root and the person morph is not justified on phonetic grounds. If the first conjugation undoubtedly uses a theme vowel in the Present Indicative, speakers could reinterpret the vowel of verbs belonging to the IIb conjugation as thematic as well.[19]

In Future and Conditional forms the situation resembles that of the Present Indicative. Verbs belonging to the IIa conjugation do not bear a theme vowel in these tenses (e.g. *perdré* 'I'll lose', *perdria* 'I'd lose'), but the lack of this vowel in IIb verbs creates phonotactic problems, and the sequences /ʃr/, /sr/, /ɲr/ and /mr/ originated by the contact of the root ending and the future/conditional marker are avoided with the insertion of a vowel:

(31)    /béns+ré/    >    [bǝnsǝré]    'I will win'

        /kunéʃ+ré/    >    [kuneʃǝré]    'I will meet'

Again, a theme vowel is used in other conjugations (e.g., *patiré* 'I'll suffer', *buscaré* 'I'll look for'), so this vowel could have been reanalyzed in IIb verbs once more, as in the other contexts mentioned above. Lloret (2002) points out that in some dialects this problematic contact between consonants is either tolerated or repaired

---

[18]Mascaró (1986) does not consider /a/ to be a theme vowel but the expression of tense. Following Oltra-Massuet (1999), I treat it instead as a theme vowel.

[19]In IIIb verbs the vowel does not correspond to the theme vowel of the conjugation, /i/, although [ǝ] could be anyway interpreted as thematic by analogy with first conjugation verbs. Interestingly, in some Catalan dialects, like Valencian, the 2nd Person singular Present Indicative of IIIb verbs contains the 'right' conjugation vowel, as in *obr*[i]*s* 'you open'. In word-final position, North-Western Catalan also provides evidence in favor of a use of a theme vowel instead of an epenthetic segment in the 3rd person singular forms of the verbs *obrir* and *omplir*, *obr*[ɛ] 'S/he opens' and *ompl*[ɛ] 'S/he fills'. An open-mid front vowel can only be underlyingly /a/ (reduced to [ɛ] in unstressed positions), the theme vowel of the first conjugation, and not the epenthetic vowel -[e]. This indicates that syllabification is constrained by morphological structure.

by consonantal epenthesis:

(32)
/béns+ɾé/ > [bən.sɾé]    (Northeastern catalan)

/konéʃ+ɾé/ > [ko.neʃ.tré]   (Valencian Catalan –Alacant–)

We could infer from the data in (32) that inserting a vowel in a theme position is something that must be avoided, at least for certain dialects. The morphophonological mismatch created by vowel epenthesis forces the use of other mechanisms that respect the correspondence between morphological and phonological structures (a consonant will never interfere with the theme position). This dissertation does not focus on verbal morphology and I will not elaborate further on this issue. However, the argument against epenthesis in theme positions is still valid for the verbal system, which implies that most vowels considered epenthetic in the literature could be reanalyzed as thematic.

Going back to the nominal system, anaptyxis is productive in loanwords only in very specific cases, with deletion being the preferred solution in illicit clusters. There have been many attempts to formalize the asymmetries found in non-native cluster resolution, focussing mainly on language transfer, syllable contact or perceptual similarity strategies (see Broselow 1992; Fleischhacker 2001, 2005; Gouskova 2002; Yun 2012, 2014, among others). Whether we adopt a view in which the native language imposes syllabic restrictions on loanwords or we prefer to rely on perceptual similarities between the source language and the incorporated word, it is obvious that not all clusters pattern in the same way, as there is a clear tendency in the languages of the world to solve rising sonority (e.g., fricative/plosive+sonorant) and falling or flat sonority clusters (e.g., sibilant+obstruent) in different ways. In Catalan, only stop + liquid/nasal clusters (= rising sonority) are split by internal epenthesis, using the same strategy as Spanish (data from Bonet 2006):

(33)
Al Sadr (Shia cleric)   >   [alsáðer]

LIDL (supermarket)   >   [líðel]

single (record)   >   [síŋgel]

In other contexts, by contrast, there is deletion of one of the segments of the cluster:

(34)    Schwarzko**pf** (brand) > [əzβáɾsko**f**]

It is important to note that none of the strategies used in nonnative cluster repairs involves word-final epenthesis (*[alsáðɾe], *[líðle], *[síŋgle]). Indeed, I am assuming for Catalan what Colina (2003) and Bonet (2006) proposed for Spanish: word-final epenthesis is not a synchronic repair mechanism of the language, although it

must have been active in previous stages. Instead, these final vowels have been reinterpreted as inflectional markers which occupy the theme position defended in this work.[20]

It is a fact that none of the new words incorporated into Catalan uses -*e* to repair word-final illicit clusters, but there are plenty of cases in the native lexicon that point in this direction. Wheeler (2005, p.258-259) provides a systematic description of the groups which require right-edge epenthesis —(35)— attributed there to the Minimum Sonority Distance requirements for Catalan syllables:[21][22]

(35)    a.   -CC of equal sonority

            fo**lr**e 'lining', corre**ct**e 'correct', pere**nn**e 'perennial'

    b.   -CC of rising sonority

            setem**br**e 'september', ama**bl**e 'kind', ri**tm**e 'rhythm'[23]

    c.   -CC of falling sonority

            a**ir**e 'air', lli**ur**e 'free', reta**ul**e 'altarpiece', diume**ng**e 'Sunday', bi**sb**e 'bishop'

Traditionally, these clusters have been analyzed as non-syllabifiable word-finally in masculine nouns and some verbal forms. These nouns have to be carefully differentiated from those where -*e* must be lexically specified, as in (36):

|  |  |  |  |
|---|---|---|---|
| (36) | *home* | [ómə] | 'man' |
| | *cotxe* | [kótʃə] | 'car' |
| | *pare* | [páɾə] | 'father' |

---

[20]In the verbal system, the final vowel of infinitives like *perdre* [péɾðre] 'to lose' is considered to be epenthetic (\**perdr*; cf. 'regular' 2nd conjugation verbs like *voler* [bolé] 'to want'). It is difficult to treat this vowel as thematic because it follows the infinitival morph /ɾ/, but even if we consider it epenthetic, its location fits with the proposal. That is, there is no inflectional position after /ɾ/, so epenthesis does not interfere with any morphological position. Alternatively, it could be considered part of the infinitival morph as well, with an allomorphic alternation -*r*/-*re*. I will not further discuss this idea as the main focus of this work is the nominal system.

[21]Codas of three or more consonants, as in *mon**str**e* 'monster', contain -CC clusters already included in (35), so I will not treat them separately.

In clusters of falling sonority in which the second element is a fricative or a plosive, the problem arises when this second consonant is underlyingly voiced, as in *diume**ng**e* 'Sunday' and *bi**sb**e* 'bishop', non when it is voiceless, cf. *Casp* (village in the Spanish region of Aragon) and *ponx* 'punch'.

[22]Majorcan Catalan permits some of these clusters only in verbal environments, as in *ompl* [ómpl] 'I fill' or *obr* [ɔ́bɾ] 'I open'.

[23]For some speakers there is total assimilation in clusters like those found in *ritme*, thus yielding an output *ri*[mm]*e*, which should therefore be grouped within (35-a). For other speakers, there is only voicing assimilation, i.e., *ri*[dm]*e*, and the cluster thus belongs to the rising sonority group.

These masculine nouns cannot have an *-e* ending due to syllabic requirements as nasals, affricates and rhotics are usual codas in Catalan (cf. *ham* [ám] 'fish hook', *raig* [rátʃ] 'ray', *rumor* [rumór] 'rumor'). If there are *-e* nouns in the language anyway, why would we need to distinguish between two *-e* groups, one with epenthesis and another one without it? It seems probable that speakers have reinterpreted the right-edge epenthetic element as an inflectional marker. As I said, epenthesis was undoubtedly an active mechanism in previous stages of Catalan, but it is not so anymore. Were epenthesis still active, the outcomes of the native lexicon and the new incorporated items would be the same —something that does not hold for Catalan (*single* [síŋgel] 'record' vs. *angle* [áŋgle] 'angle').[24] Therefore, the structures for both groups, depicted in (37) (syllabification problem) and (38) (no syllabification problem), do not differ and must have a close-mid front vowel in their lexical entry (in Pallarese, a schwa in Central Catalan), which occupies the theme position:
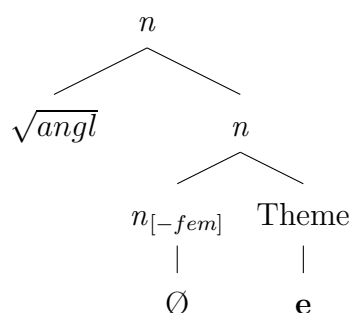
(37)     *angle* 'angle'                    (38)     *pare* 'father'
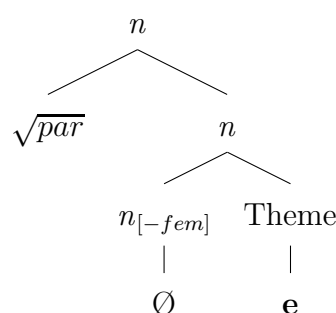


It has been argued in section 4.1.1 that in Pallarese nouns the featural exponents corresponding to gender are transferred to the theme position, which already contains a V-slot where those features are realized (as in (39)), yielding a structure such as that of (40) (these structures correspond to examples (14) and (15) in 4.1.1 but are repeated here for the purpose of clarity):

---

[24]It is true that perceivable excrescent vowels in the source language could affect the repair strategy. Nevertheless, whereas other similar languages, like Portuguese, still use word-final epenthesis for loanword adaptation (English to Brazilian Portuguese: blog > [blɔg**i**], Xavier 2013), Catalan always prefers deletion or anaptyxis.

(39)

```
              n
            /   \
        √kaz      n
                /   \
    n[+fem]↔[+low]   Theme
                       |
                       V
```

(40)

```
              n
            /   \
        √kaz      n
                /   \
         n[+fem]    Theme
            |         |
            Ø         a
```
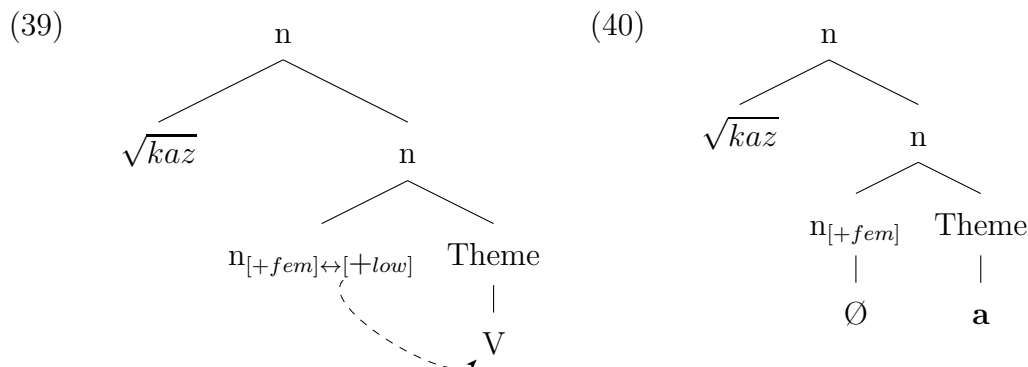
As already mentioned in this chapter, it must be clearly stated that the derivation in (39) and (40) is only valid for regular unmarked cases (those with a -Ø ending in the masculine and -a in the feminine). The theme position in (39) will only contain an underspecified vowel in regular cases (before it gets fully specified with the place features associated with gender, as in (40), which are the vast majority of nouns in Catalan, and parallels the scenario found in Pallarese masculine and feminine clitics analyzed in 5.1.1 (see that section for the details of the OT analysis).[25] In all other cases, a vowel still occupies the theme position, but it is already contained in the lexical entry. In other words, marked nouns are lexically listed as (37) and (38).

We can conclude, then, that the theme position constrains the location of epenthesis and forces an interpretation by which any right-edge vowel of the word is considered thematic. This is the case in words like *angle* 'angle' (that we have just seen), which underwent epenthesis in previous stages of the language to repair an illicit coda cluster, but, synchronically, their final vowels are interpreted as theme vowels due to their position within the word (meaning there is no epenthetic process at all).

Besides Catalan, right-edge epenthesis also seems to be also avoided in other Romance languages. Repetti (2012, p. 180) claims something similar for Italian, pointing out that "word-final position is a morphologically-salient position in Romance languages" and, therefore, epenthesis tries not to interfere with this position. As a result, instead of inserting excrescent or epenthetic vowels in the codas of loanwords (the focus of her analysis), the use of -*o*, -*a* or -*e*, all of them inflectional markers of Italian, "allow<s> the borrowed noun (or adjective) to be integrated into a declension class" (Repetti, 2012, p.183).[26] This phenomenon is labeled 'mor-

---

[25]See also 7.1 for a detailed analysis of masculine nouns like *angle* 'angle', whose final vowel is not considered epenthetic in my proposal. Traditional studies of Catalan assume, instead, that it is in fact epenthetic, and confront nouns like *angle* with those like *pare* 'father', with a vowel that must be necessarily specified in the lexical entry. I conflate these two groups into one and assume that the vowel is always specified in the lexical entry.

[26]Repetti (2012) documents [ə],[e] and [ə] as paragogic vowels in certain loanwords. Interestingly, in the experiment conducted to test the quality of the vowels, none of them was identical to unstressed lexical -*e*, the inflectional marker. This clearly indicates that the status of epenthetic

phological epenthesis' in Cardinaletti & Repetti 2008, and it is in fact the same idea developed previously in Lloret & Viaplana (1992), who name it *epèntesi morfològica* (literally, morphological epenthesis). The idea is simple: instead of epenthetic vowels, languages can use inflectional markers to repair syllable structure. This is why we find nouns like (41) in North American varieties of Italian (Repetti, 2006) or place names and borrowings like (42) in Standard Italian, which do not insert the expected default epenthetic vowel:[27]

|      |          |   |             |      |           |   |                   |
|------|----------|---|-------------|------|-----------|---|-------------------|
|      | rug      | > | [rágg**a**] |      | Zurich    | > | *Zurig*[**o**]    |
| (41) | brick    | > | [bríkk**o**]| (42) | Stockholm | > | *Stoccolm*[**a**] |
|      | business | > | [bisníss**e**] |   | beef steak| > | *bistecc*[**a**]  |

The morphological structure of the nominal system imposes certain restrictions on loanword incorporation. Declension markers are preferred as repair strategies because they ease the classification of recently incorporated words, and *-i*, the default epenthetic vowel of Italian, is avoided. Actually, Repetti (2012) claims that *-i* is only found in word-initial or word-internal position –(43)– because at the right edge of the word it becomes a marked vowel that signals plurality in both masculine (*cap**i*** 'heads') and feminine nouns (*chiav**i*** 'keys').

|      |            |   |                    |
|------|------------|---|--------------------|
| (43) | psicologo  | > | [p**i**sicólogo]   |
|      | in Svizzera| > | in [**i**]Svizzera (formal contexts) |

It is important to take into account that this is not exclusive to Italian. In Spanish, this kind of 'declension-class-incorporation'-like epenthesis is present in some varieties:

|      |                          |           |   |              |                 |
|------|--------------------------|-----------|---|--------------|-----------------|
| (44) | Latin American Spanish:  | champagne | > | champañ[a]   | (*champa[ɲ])    |
|      | Peninsular Spanish:      | champagne | > | champá[nØ]   | (*champa[ɲ])    |

Instead of resorting to the regular epenthetic vowel *-e*, we see in (44) how Latin American Spanish uses a word marker that allows the French borrowing *champagne* 'champagne' to be treated as a regular feminine *-a* noun. In other varieties, like Peninsular Spanish, it is preferable to depalatalize the nasal consonant to make it alveolar (or velar depending on the dialect) and thus become a licit coda. Feature changing or feature deletion is, in fact, widely used in Peninsular Spanish, a dialect that, as already mentioned, has completely abandoned word-final epenthesis

---

vowels is different from that of the vowels of the native lexicon.

[27]The vowel *-e* of [bisníss**e**] is a word marker, as in *can**e*** 'dog', not an epenthetic vowel.

as a repair mechanism (e.g., 'beef steak', incorporated into Italian as *bistecc**a***, is pronouned in Peninsular Spanish as [bisté] or [bistél] –liquids are tolerated in word-final position–).

As for Catalan, this process of morphological epenthesis is similar to the already mentioned cases of OCP-sibilant plurals like *goss**os*** 'dogs', built upon the base *gos* 'dog'. The appearance of the masculine morph *-o* in the plural is phonologically conditioned but imposed by the morphological structure, and thus preferred over the default epenthetic vowel of Catalan. The fact that, under my analysis, the phonological features associated with gender are floating allows their association with the theme position under certain phonological conditions. This kind of plural is analyzed in chapter 7, whereas the similar alternation between *-Ø* and *-o* in 3rd person singular masculine clitics/definite article (*l/lo*) is dealt with in 5.1.1.

All in all, these data make clear that epenthesis is sensitive to morphological structure, at least in certain languages. It may not be the case during the initial stages of the epenthetic process in which the phonological problem is originated, but later on the noun system forces a reanalysis to incorporate the epenthesized words into regular declension classes.

## 4.2   Gender in Pallarese Catalan

Before we move on to the OT analysis of gender inflection in Pallarese Catalan clitics in chapter 5, we need to understand how the nominal system works. Pallarese Catalan, like Central Catalan, has a binary system for gender (masculine and feminine) and number (singular and plural). The distribution of the inflectional endings in the singular is presented in table 4.1:[28]

| | Masculine | | Feminine | |
|---|---|---|---|---|
| [Ø] | *cor*[Ø] | 'heart' | *sal*[Ø] | 'salt' |
| [o] | *llor*[o] | 'parrot' | mot[o] | 'motorcycle' |
| [a] | *di*[a] | 'day' | *cas*[a] | 'house' |
| [e] | *par*[e] | 'father' | *class*[e] | 'class' |

Table 4.1: *Pallarese inflectional endings*

The vast majority of masculine nouns have a *-Ø* ending, while the feminine is

---

[28]Other endings, such as *-us* (*cactus* 'cactus') or *-is* (*brindis* 'toast') or *-i* (*bigoti* 'moustache'), are found in the language but are rather marginal. See 3 and 7.1 for a more detailed analysis of the different inflectional exponents of Catalan.

realized as -*a*.[29] The differences in vowel reduction with regard to Central Catalan provide a slightly modified inventory of inflectional endings. Whereas the -*o* ending is realized as -[u] in Central Catalan, it surfaces as -[o] in Pallarese. In addition, the -*e* and -*a* inflectional markers are conflated into phonetic -[ə] in Central Catalan, but they are realized separately as -[e] and -[a] in Pallarese.

Although historical apocope is behind the absence of an overt exponent in the masculine, there is still a group of nouns, like *lloro* 'parrot', which end in -*o*, whose inflectional marker must be specified in the lexical entry. This marked -*o* ending (which is relatively frequent) is also used, as already mentioned, in plurals ending in a sibilant in contexts where the contact with the plural morph /z/ creates a phonotactic problem, as shown below:

(45)

| | | | |
|---|---|---|---|
| *gat* 'cat' | > | *gats* 'cats' | |
| *lloro* 'parrot' | > | *lloros* 'parrots' | |
| *pas* 'step' | > | *pass**o**s* 'weights' (**pass*) | |

In a word like *passos*, the use of a marked inflectional ending -*o* prevents the contact of two sibilants. A phonological problem is thus repaired via morphology, whereas a purely phonological solution that would imply the use of the default epenthetic vowel (-[e] in Pallarese, schwa in Central Catalan) is discarded. This phenomenon has been labeled *epèntesi morfològica* 'morphological epenthesis' in Lloret & Viaplana (1992), a term also used in Cardinaletti & Repetti (2008)'s analysis of Northern Italian dialect clitics (as we just saw in 4.1.2). Following this idea of morphological epenthesis, what I aim to demonstrate in this dissertation is that the morphological structure of Catalan nouns conditions the location of epenthesis and forces the use of -*o* in this set of plurals (as well as in other contexts).

Several works have analyzed the nature of this vowel (Mascaró 1985, 1986; Lloret & Viaplana 1992; Wheeler 2005, among others), but it is in Bonet et al. (2007) that the phenomenon is fully formalized. The analysis of the Central Catalan nominal system provided therein resorts to allomorphy to account for the insertion of -*o* in contexts where syllabification is not possible with a -Ø ending in the plural, as in (45). Their proposal has strong implications for the overall inflectional nominal system of Catalan and is the topic of the next section, 4.2.1, as in chapter 5 I present

---

[29]In Pallarese, the feminine marker is phonetically -[a] in the singular, e.g. *cas*[a] 'house', but -[e] in the plural, e.g., *cas*[e]*s* 'houses'. Jiménez (2002) and Massanell (2011) propose that, for Northwestern Catalan (the dialectological group Pallarese belongs to), the feminine marker is underlyingly /a/, but due to a phonological process it is pronounced as -[e] in unstressed closed syllables in final position (after the addition of the plural morph -/z/ to the stem, in this case). The same is assumed here for Pallarese.

an alternative analysis that does not resort to allomorphy.[30]

## 4.2.1   Allomorphy?

The main idea behind Bonet et al. (2007)'s analysis is that allomorph selection is always controlled by the phonological component. That is, in an OT grammar, constraint interaction is responsible for the choice of the surface allomorph not only in phonologically conditioned environments but in all cases.

In their view, each lexical entry contains all allomorphs stored in a totally ordered set relying on markedness. Markedness is determined on a frequency basis, as in table 4.2 (the vowel quality of the inflectional endings is not the same as in Pallarese due to a different process of vowel reduction in Central Catalan, the dialect studied in Bonet et al. 2007):

|            | *Unmarked* | *Marked* |
|------------|------------|----------|
| *Feminine*  | ə          | Ø        |
| *Masculine* | Ø          | u > ə    |

Table 4.2: *Markedness in inflectional endings (Central Catalan)*

Given this assumption, the lexical entries for gender look as follows:

(46)
$$\text{Masculine: } \{Ø > u > ə\} \quad \text{Feminine: } \{ə > Ø\}$$
$$\text{Full lexical entry:} \quad /\text{pás}\{Ø > u > ə\}/ \text{ 'step'} \quad /\text{ták}\{ə > Ø\}/ \text{ 'stain'}$$

The constraint that guarantees the surfacing of unmarked allomorphs, -Ø for the masculine (*pas* 'step') and -*a* for the feminine (*taca* 'stain'), is PRIORITY (Mascaró, 2007). PRIORITY favors faithfulness to the first allomorph in the hierarchy shown in (46).

(47)   PRIORITY: Respect lexical priority (ordering) of allomorphs.

Although the vast majority of nouns take the default allomorph for both masculine and feminine, the other marked cases (like *lloro* 'parrot' in the masculine or *sal* 'salt' in the feminine) still need to be accounted for, as PRIORITY on its own would always generate regular nouns (i.e., with a -Ø ending in the masculine and -*a* in the feminine). To prevent this, a subcategorization requirement in the input is needed in order to generate the correct output. Marked cases are thus lexically specified

---

[30]Bonet et al. (2007)'s proposal has already been reviewed in section 3.1.1, but I repeat the general ideas in the next section to better highlight the differences with regard to my proposal.

with a diacritic (e.g., *lloro* 'parrot': /ʎɔɾ$_u$/), and the constraint Respect ensures that the subcategorization requirement is satisfied:

(48)     Respect: respect idiosyncratic lexical specifications of allomorphs.

Thus, the interaction between Respect and Priority yields both regular and marked cases, e.g. *pas*[Ø] and *llor*[u], respectively, given that Respect always dominates Priority.

A different scenario is found in the above-mentioned masculine plurals with morphological epenthesis (*pas* > *pass**o**s*). When there is no phonological conflict or lexical specification, the default allomorph -Ø is selected (*pas* 'step'), but the interplay between faithfulness constraints such as Priority and phonotactic restrictions against a sequence of sibilants (OCP-sibilant, *\*pass*) generates an output like *pass*[**u**]*s* 'steps', where the use of a marked masculine allomorph (i.e., /{Ø > **u** > ə}/) is preferred over a (phonological) epenthetic solution.

Although the use of allomorphy in combination with Priority seems appealing for the morphological epenthesis cases, chapter 5 will argue against this view, as, at least for Pallarese Catalan, a simpler solution can be offered in autosegmental terms.[31] There is no need to resort to allomorphy if we assume that the exponents for gender values are floating features needing only to be linked to timing units (and, therefore, pronounced) under certain conditions, as already shown in previous sections. With this move, we can dispense with a hierarchically ordered rich input for inflection, which increases complexity in the generative system.

In fact, this idea of floating features as a rescue strategy developed in chapter 5 for Pallarse clitics is similar to del Gobbo (2001)'s analysis of Italian definite and indefinite articles. The examples in (49), (50), (51) and (52) illustrate the phonological conditioning of the article in Italian:[32]

(49)     *Masculine definite article*

     a.   *il ponte*     'the bridge'
     b.   *lo specchio*  'the mirror'
     c.   *l'albero*    'the tree'

(50)     *Masculine indefinite article*

---

[31]The same idea could be extended to Central Catalan and other dialects (see section 8.2), although the main focus of this thesis is Pallarese.

[32]I do not provide here a complete review of Italian articles. Only the examples that are useful to understand the details of the proposal in chapter 5 are shown.

    a.   *un ponte*    'a bridge'

    b.   *uno specchio*  'a mirror'

    c.   *un albero*    'a tree'

(51)    *Feminine definite article*

    a.   *la casa*    'the house'

    b.   *l'amica*    'the friend'

(52)    *Feminine indefinite article*

    a.   *una ragazza*    'a girl'

    b.   *un' epoca*    'an age'

For the masculine definite article, illustrated in (49), the form [l] appears before vowels, [il] before consonants (and tolerated clusters) and [lo] before illicit clusters, whereas its indefinite counterpart –(50)– has only two forms, [un] before consonants, vowels or some clusters and [uno] before illicit clusters. As for the feminine definite —(51)— and indefinite article —(52)—, [l] and [un] occur before vowels and [la] and [una] in all other contexts.

    The lexical entries given in del Gobbo (2001) to explain the distribution of Italian articles are the following (the lexical entries of the other articles that are irrelevant for the present purposes are obviated):

(53)
    a.   Definite masculine singular:    /il, l(o)/

    b.   Definite feminine singular:    /la, l(o)/    (*l(o)* = masculine)

    c.   Indefinite masculine singular:  /un(o)/

    d.   Indefinite feminine singular:  /una, un(o)/   (*un(o)* = masculine)

The general assumption in Del Gobbo's proposal is that, in Italian, there are floating -*o* segments (indicated in parentheses) in the masculine definite and indefinite singular articles that are inserted to help syllabify the first element of complex onsets in nouns as codas (e.g., [los.pék.kjo], *lo specchio* 'the mirror'). Besides, the absence of the gender marker in feminine singular prevocalic contexts, as in *un'epoca* 'an age' (**una epoca*), is not due to apocope; *un* is, instead, the masculine allomorph, which avoids the hiatus that would originate the use of the feminine indefinite article.[33]

    The analysis of Pallarese Catalan presented in this dissertation, although rely-

---

[33]In Steriade (1999)'s terms, this is a case of Lexical Conservatism, similar to the *beau/bel* alternation of French adjectival liaison.

ing on floating features for syllabification purposes as in Italian, does not need to stipulate allomorphs for certain cases and floating vowels for others, as del Gobbo does for Italian. All the differences in Pallarese definite articles and 3rd person pronominal clitics can be accounted for by resorting only to the theme position where gender is realized. The features provided by the morphological component are always floating (for both masculine and feminine) and they will be linked to a structural position (i.e., the V-slot of the theme position) depending on the constraint ranking, which generates all the different phonetic forms. Furthermore, in del Gobbo (2001)'s proposal, the lexical entries for the feminine definite and indefinite articles contain masculine and feminine exponents, as shown in (53), and a constraint GENDER ("use the morphologically correct article"), intertwined with phonological restrictions, decides upon the final output. In my analysis of Pallarese, the differences between definite articles (*lo/l* in the masculine and *la/l* in the feminine) are easily captured by the interaction of phonological constraints, without the need to invoke allomorphy or morphological constraints. This is crucial in the proposal, as morphology and phonology are strictly delimited in the tasks they have to accomplish. Any interaction between both modules increases complexity in the overall computational system, and thus it is preferrable to maintain a phonological component with constraints that refer exclusively (when possible) to phonological objects. A constraint like GENDER is a descriptively adequate device to generate the expected outputs, but gender values constitute morphosyntactic information, and therefore phonology should not have access to the "correct" gender of an article (in del Gobbo's terms). As defended in this dissertation, morphology is in charge of translating the output of the syntactic computation into phonological material, but mixing constraints like GENDER with NOHIATUS (purely phonological, also used in del Gobbo's analysis) in the phonological module assumes that phonology is not blind to syntax. As argued in van Oostendorp (2015) phonology and syntax cannot read each other's features; only structural information, i.e., headedness and boundaries, can be interpreted between them, and therefore GENDER-like constraints should be avoided in an OT phonological grammar. Finally, the assumption of a theme position correctly predicts the insertion of the masculine morph -*o* instead of an epenthetic vowel in masculine clitics, something that could not be grounded in the morphological structure of the clitic in del Gobbo (2001):

> "(...) In fact, we might wonder why in this case /o/ is the chosen vowel for epenthesis, and not /e/, which seems to be the usual epenthetic vowel for loanwords in Italian. We would have in this case a morphological motivation to predict the quality of the inserted vowel: /o/ is chosen

because it usually marks the masculine gender and singular number for
Italian nouns. Since full epenthesis is usually phonologically driven, we
would have to accept that here it is morphologically determined. Because
of this reason, in this paper I will adopt the floating vowel hypothesis,
leaving a more adequate answer to the issue to further developments."
(del Gobbo, 2001, p. 8)

It is thus clear that the insertion of gender exponents (i.e., floating phonological
features under the present analysis) is constrained by morphological structure, but
just by delimiting the tasks that phonology and morphology have to accomplish we
can get rid of gender allomorphy, at least for Pallarese Catalan.

Before proceeding to the morphophonological analysis of gender inflection in
Pallarese clitics, some syntactic approaches devoted to gender are reviewed in section
4.2.2 in order to properly grasp the proposal of gender realization in the theme
position in chapter 5.

### 4.2.2 Syntactic approaches to Gender

Although there are many morphophonological analyses of Romance nominal inflec-
tional systems in the literature (see 3.1 for a review of Catalan and Spanish pro-
posals), the analysis of gender from a purely syntactic point of view has received
relatively little attention. That is, even if the role gender plays in agreement sys-
tems has been well studied in the field of generative linguistics, its *location* within
the syntactic structure has not been clearly defined in most cases.[34]

It is clear that gender controls agreement in the nominal domain, as shown below
for Italian:

(54)    Quell**e**        ginocchi**a**     sono     grand**i**
        that.FEM.PL knee.FEM.PL be.3PL big.FEM.PL
        'Those knees are big'

The default feminine plural exponent in Italian is *-e* (which spells out both gender
and number in one single morph), but (54) shows how other inflectional markers like
*-i* and *-a* can appear in feminine plural contexts. The feminine plural word *ginocchia*
determines the gender of the demonstrative and the adjective, but none of the words
in (54) share the same exponent, showing a non-direct correlation between gender

---

[34]As I said in chapter 3, I mainly analyze 'formal' gender in this dissertation, i.e., the arbitrary
association of a morphosyntactic feature —[±fem] in the case of Catalan— to nouns. There is also
a link between this formal expression of gender and biological sex (usually –although, crucially,
not always–, biological feminine referents are assigned feminine morphology, and viceversa for
masculine nouns), but I do not treat these issues in depth here.

and inflection (something which has already been shown for Catalan and Spanish in chapter 3). Moreovoer, the singular counterpart for *ginnochia* is *ginocchio*, a masculine noun, as illustrated in (55):

(55)    Quel**Ø**      ginocchi**o**   è        grand**e**
        that.MASC.S knee.MASC.S be.3PL big.MASC.S
        'That knees is big'

The example in (55) shows, again, that there is not a one-to-one correspondence between inflectional markers and gender, as the default exponent for masculine singular, -*o*, appears only in the noun, not in the demonstrative or the adjective, which are also masculine —cf. (54). In addition, the data in (54) and (55), where gender changes in the singular and the plural of the same lexical entry, raises another question: where is gender stored, and how is it related to noun structure?

The answer to this puzzle is still an ongoing debate in the literature within the framework assumed here but, basically, there are three options with regard to the placement of gender in nouns:

- Gender is an inherent property of roots

- Gender has an independent syntactic projection

- Gender is stored on little *n*

As all the options mentioned above face some problems, they will be discussed separately in 4.2.2.1, 4.2.2.2 and 4.2.2.3.

### 4.2.2.1   Gender as an inherent property of roots

Gender could be assumed to be a property of roots. This assertion seems *a priori* adequate from a descriptive point of view, but a closer look at the nominal system shows that the picture is far more complicated. Take for instance the Catalan examples in (56) and (57):

(56)    a.    *nena* 'girl'
        b.    *nen* 'boy'

(57)    a.    *porc* 'pig'
        b.    *truja* 'sow'

In (56) we have a masculine/feminine pair with the same phonological spell-out, the only difference being the inflectional marker that the two nouns bear. A masculine/feminine pair is also found in (57), but it does not share the same set of

exponents. Two different outputs must, presumably, be stored for nouns such as in (57) which differ only in their biological sex but are otherwise semantically equivalent.. Therefore, it could be argued that the same applies to nouns like those of (56), even if it seems redundant to store the same phonological information twice, as illustrated in (58):[35]

(58)  a.  $\sqrt{\text{NEN}}_{[+fem]} \leftrightarrow$ /nɛn/
      b.  $\sqrt{\text{NEN}}_{[-fem]} \leftrightarrow$ /nɛn/

The analysis of same-root nominals (in Kramer 2009's terminology) faces the same problem, as well. These nouns, depicted in (59), have exactly the same phonological shape whether their gender is feminine or masculine, something that depends on the biological sex of the entity they refer to:[36]

(59)  a.  *un/una cantant* 'a singer' (m./f.)
      b.  *una/una testimoni* 'a witness' (m./f.)

Note that the nouns in (59) also go against "simple" analyses of gender that consider this feature to be lexically listed for each noun (for the noun as a whole, not specifically in the root), which can be appealling at first glance. That is, nothing distinguishes phonologically masculine/feminine pairs like *cantant* or *testimoni*, and "it clearly would be undesirable to have two homophonous, synonymous nouns for each of these cases, one with masculine gender and one with feminine gender" (Kramer, 2015, p. 27)

The masculine and feminine forms shown in the Catalan examples above are non-distinguishable from a phonological point of view (in (56) it is only the inflectional ending which indicates the gender of the noun, and in (59) not even that), and, thus, having two different roots for the same noun is not a straightforward solution. As Kramer (2009, p.120) points out, "(...) this might be acceptable if only a small number of nominals were ambiguous, but ambiguity is in fact quite common". Although the language analyzed in Kramer (2009) is Amharic, the same conclusion can be drawn for Catalan, as nouns like those of (56) and (59) do not constitute

---

[35]Gender is phonologically spelled out separately, e.g., [+fem] ↔ /a/, and thus the final output for the feminine noun is /nɛna/. Nonetheless, the root has the same realization whether it is marked as [+fem] or [–fem], as shown in (58).
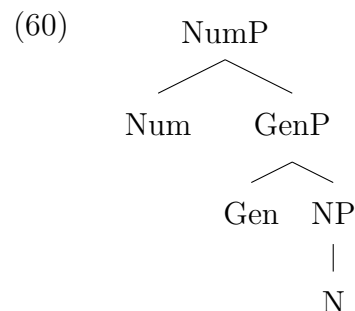
[36]Amharic nouns are not inflected for gender (gender is instead indicated by agreement with definite markers, demonstratives or verbs), and thus Kramer labels same-root nominals cases like (59), which contrast with the examples in (57) that have two different roots for masculine and feminine referents. In Catalan, both (56) and (59) must be considered same-root nominals, although overt inflection distinguishes the masculine from the feminine only in (56).

isolated cases in the language and are instead very high in frequency.[37]

In addition, the fact that some derivational suffixes can determine the gender of a noun, e.g, *-esa* creates feminine nouns (as in *bellesa* 'beauty') but *-isme* creates masculine nouns (as in *atletisme* 'athletism'), goes also against the representation of gender in roots (in these cases, gender features must be associated with suffixes; see 7.1.1). Therefore, treating gender as an idiosyncratic property of roots does not seem the most convincing alternative.

### 4.2.2.2  Gender as a separate syntactic projection

A different line of thought within generative linguistics advocates an independent syntactic projection where gender is spelled out, namely GenP. This is the proposal advanced for the nominal domain in Picallo (1991), whose claim is that GenP is dominated by NumP (assuming a complex structure of the DP), as shown in (60):

(60)

```
            NumP
           /    \
        Num     GenP
               /    \
            Gen      NP
                      |
                      N
```

Successive cyclic movement raises the N head to NumP, and therefore the syntactic tree in (60) reflects the surface order of morphemes, with number being the outermost feature of the noun. Formal gender, however, is an unintepretable feature with no semantic content ("we will propose that Gender projects onto a syntactic functional category in all Catalan nominals, even when it does not have semantic content", Picallo 1991, p.282), and as Kramer (2009) points out, GenP contradicts Chomsky (1995)'s idea against functional projections that only contain uninterpretable features.

Picallo (2006, 2008) proposes a more refined structure to accommodate gender in the nominal domain (as well as classifiers in Bantu languages), arguing for an interpretable feature [CLASS] that depends on the functional category *c* where gender —which is uninterpretable and a property of the N head— can be valued. The problem with this idea is that it is not clear at all why [CLASS] should be considered

---

[37]It is important to highlight that it is irrelevant if we are assuming a lexical-realizational model of morphology or not, as the problem is met in any framework. Having two different roots with the same set of exponents does not seem an optimal solution for the system.

an interpretable feature; according to Picallo, this has to do with a process of "entity categorization":

> "(...) I have suggested that gender declension should be associated with a functional category *c* that contains an *abstract interpretive feature* that relates grammar with processes of *entity categorization.* The valuation of the feature hosted in this syntactic category can crosslinguistically surface in several ways: as an independent lexeme (i.e. a noun classifier) or, more abstractly, as formal gender declension." (emphasis mine, Picallo 2006, p. 111)
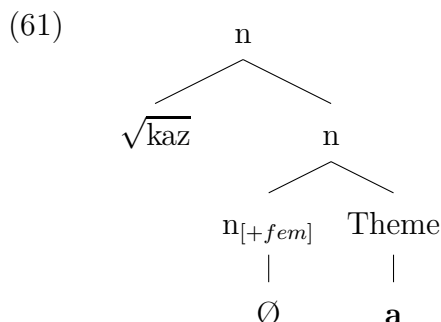
However, as already mentioned, gender —when it is not related to sex— does not contribute to the semantics of the noun.[38] In this scenario, it is not clear how the feature [CLASS] associated with entity categorization can be interpretable if it does not add any semantic value to the noun.[39] The solution found in Picallo works from a theoretical point of view, but it does not seem empirically adequate and, therefore, a proposal that better captures the arbitrary use of formal gender must be preferred.

### 4.2.2.3  Gender on little *n*

It has been assumed throughout this chapter that the phonological spell-out of morphosyntactic gender, which is attached to little *n*, is transferred to the theme position in the nominal domain. Thus, a feminine word like *casa* 'house' has the following —final— phonological spell-out (after the intermediate step where the floating features associated with [+fem] have been attached to the V-slot of the theme position):

---

[38]Leaving aside the morphological expression of natural (sexual) gender distinctions, one can say that the interpretation of common nouns, either mass or count, is not affected by their belonging to a particular gender type (Picallo, 2008, p. 50)

[39]See Arsenijević (2009) for an analysis of gender as a grammaticalized classifier system in Serbo-Croatian, which has some similarities with Picallo's proposal. However, neuter nouns in Serbo-Croatian are determined on some semantic basis, as the author claims that neuter count nouns fail to formally express uniform atomicity. The same does not hold for Catalan, and thus gender cannot be treated equally.

(61)

```
              n
           ╱     ╲
       √kaz        n
                 ╱   ╲
          n_[+fem]   Theme
             |         |
             Ø         a
```

The *n* head, and not the root, is specified for gender, as shown in (61). This is a proposal that has been made for Romance and other languages, but, like the other two alternatives already reviewed, it also faces some drawbacks.

Acquaviva (2009), among others, provides an analysis of Italian nouns where roots are radically underspecified. He shows that Italian plural doublets like *ossa/ossi*, depicted in (62), are better analyzed under the assumption that nouns can be assigned masculine or feminine gender only under certain licensing conditions.
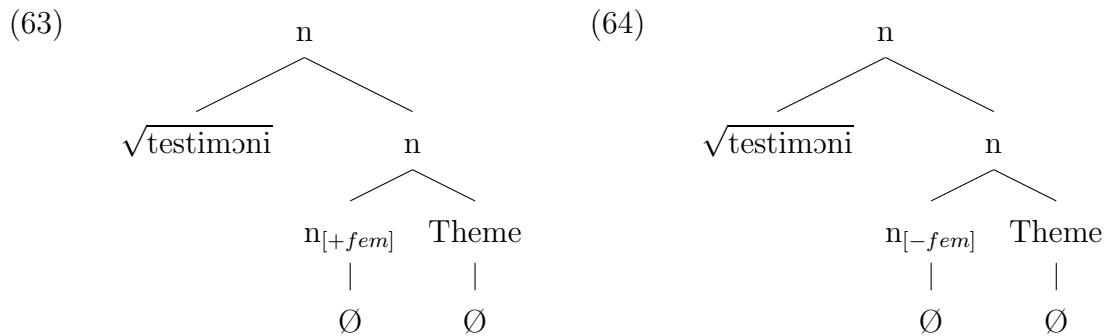
(62)  a.  *osso*   (m. s.)    'bone'
      b.  *ossi*   (m. pl.)   'bones'   (unconnected wholes)
      c.  *oss**a**   **(f. pl)**    'bones'   (connected parts)

That is, if gender is a feature of the root, we must assume that we have two different roots for the examples above, which seems undesirable. Instead, if gender is a property of the *construction*, we can solve this problem by assuming that the interpretation in c. is only licensed when the root is merged onto a feminine *n* head ($n_{[+fem]}$). This is a general property of the language, and, thus, all feminine words are licensed *only* under $n_{[+fem]}$, whereas masculine words are licensed *only* under $n_{[-fem]}$. Therefore, as Kayne (2006, p. 290) points out, the different *n* heads are a sort of 'list' of the roots they can be merged with:

> "From this perspective, feminine gender can be merged with a noun only if (in the unpredictable cases) that (suffixal) noun is contained in its list, in essence a selectional property of feminine gender (feminine gender via agreement, on adjectives, determiners, etc. will involve something more)."

Kramer (2009, 2014, 2015), as already shown in 4.2.2.1, argues against the "undesirable repetitions" of gender being lexically listed for every N, and assumes as well that gender is a feature on the nominalizing head *n* ("this would mean that assigning gender to a root plays an essential part in turning that root into a nominal, which seems intuitively correct" (Kramer 2015, p. 33). As in Acquaviva (2009), the

use of licensing conditions on roots allows her to better capture the distribution of gender in Amharic (the core of her studies). This choice is particularly adequate for the analysis of same-root nominals in Amharic, quite common in the language, and is also useful for Catalan, e.g., *un/una testimoni* 'a witness' (m./f.) in (63) and (64). If the gender feature is on little *n*, it is clear that we do not need to duplicate roots because it is the *n* head that changes, depending on the referent:[40]

(63)

$$
\begin{array}{c}
n \\
\diagup \quad \diagdown \\
\sqrt{\text{testim}ɔ\text{ni}} \qquad n \\
\diagup \quad \diagdown \\
n_{[+fem]} \quad \text{Theme} \\
| \qquad | \\
\varnothing \qquad \varnothing
\end{array}
$$

(64)

$$
\begin{array}{c}
n \\
\diagup \quad \diagdown \\
\sqrt{\text{testim}ɔ\text{ni}} \qquad n \\
\diagup \quad \diagdown \\
n_{[-fem]} \quad \text{Theme} \\
| \qquad | \\
\varnothing \qquad \varnothing
\end{array}
$$

As illustrated in (63) and (64) for Catalan, the use of different *n* heads allows a simple treatment of same-root nominals: we can still maintain a bare root because it is the construction which determines whether the noun is masculine or feminine.[41] For a language like Amharic, which has a formal gender system that depends more on biological sex than Catalan, the licensing conditions imposed by different *n* heads (different from those of Catalan) can deal with all the idiosyncrasies of the language, something that a gender-on-root analysis is not able to do.

Amharic uses a binary system for gender with masculine and feminine nouns. Animate nouns can have different roots for male/female pairs, like *bal* 'husband' - *mist* 'wife', but the vast majority use the same root for both genders, as in *hakim* '(female/male) doctor'. The default gender value is masculine (if the gender of the referent is not known, then it is used as a masculine for agreement purposes), but some animals are exceptionally feminine by default, as shown in (65):

(65)  ayt'  -wa
      mouse DEF.FEM
      'the mouse'

---

[40]I use *testimoni* in the examples, but the same that holds for *testimoni* can be applied to pairs like *nen-nena* 'boy-girl', since there is also no need to duplicate roots in these nouns (the difference lies on the inflectional ending, which is the phonological spell-out of [±fem]).

[41]This is the final spell-out of a feminine/masculine pair like *testimoni*, but we will see in 7.1 how the the realization process of the theme vowel is different in both cases; it is regular for the masculine noun but, in the feminine, as the theme vowel is not the default (i.e., *-a*) the position contains a zero exponent.

However, if the biological gender of the animal is known, it overrides the formal
feminine default gender, as in (66):

(66)    ayt'    -u
        mouse DEF.MASC
        'the male mouse'

This leads Kramer to conclude that natural gender (if known) determines the formal
expression of gender in animate nouns. As for inanimate nominals, almost all of them
are masculine (only 20-30 are feminine under Kramer's estimations).

     Given this distribution of gender in the nominal system, Kramer proposes four
types of $n$ heads that can generate all the nouns of the language:

(67)     *Types of* n *in Amharic*

   a.   n   $_i$[+fem]    Female natural gender

   b.   n   $_i$[–fem]    Male natural gender

   c.   n                 No natural gender (or natural gender irrelevant/unknown)

   d.   n   $_u$[+fem]    Feminine grammatical gender

     In (67), it is shown that some $n$ heads are specified with feminine and masculine
features, which can also be interpretabele or uninintepretable (the latter only for
the feminine). If the referents are animate, it is clear that they are licensed under $n$
heads which bear interpretable $_i$[+fem] or $_i$[–fem] features for feminine and masculine
nouns, respectively. In addition, positing a bare $n$ head, with no gender specification,
also allows a straightforward generation of same-root nominals where the sex of the
referent is not known, as it will be realized with masculine agreement since masculine
is the default (but it is not specified on $n$ because it is used as a neuter). Finally,
the uninterpretable [+fem] feature is used for (rare) inanimate feminininе nouns and
animate nominals with feminininе default gender alike. Therefore, the exceptional
animals that are feminine in Amharic, like *ayt'* 'mouse' (gender unknown), will
only be licensed under $n_u$[+fem], not under bare $n$ (which would trigger default
masculine agreement). If the gender is known, though, these nouns will be licensed
under $n_i$[+fem] or $n_i$[–fem].

     Kramer's proposal is adequate both theory-internally and empirically as it cap-
tures all the facts about Amharic gender inflection but, nevertheless, any analysis
based on licensing conditions faces one basic problem: what determines that a noun
can be merged onto $n_{[+fem]}$ or $n_{[-fem]}$ (or other classes of $n$ heads as in Amharic)?
That is, why is it that certains roots can only be licensed under masculine $n$ heads

and not under feminine $n$ heads, and vice versa? Some kind of information needs to be marked somehow in order to get the correct gender and, therefore, the argument for gender features on little $n$ becomes somehow circular. This problem is acknowledged by Acquaviva (2009, p. 6):

> "To say "a noun has gender X", for instance, means in this perspective "a root Vocabulary item is licensed in the context of [n] with gender X". This type of licensing is clearly linguistic knowledge, but it is not deterministically grammar-driven. Licensing statements that apply to lists of roots, by themselves, are not more (nor less) arbitrary than explicit specifications on each root; the difference is empirical, and I will argue that viewing *lexical information as a property of constructs* rather than roots is empirically more successful. *The crucial difference from earlier approaches is that meaning arises in a construction, not in a root.*" (emphasis mine, Acquaviva 2009, p. 6)

As Acquaviva points out, licensing statements consist of lists of roots under a particular $n$ head, which is as arbitrary as specifiying roots. However, as he emphasizes, it is the construction itself which determines the gender of the word. This has one clear advantage: it avoids the duplication of roots (and it is therefore preferable). We can then solve the puzzle posed by the Italian examples contrasting gender in the singular vs. plural in (54)-(55) and in (62). Moreover, having 'pure' bare roots is also more adequate to capture the distribution of morphosyntactic categories in languages. Nouns become nouns when they are merged to little $n$, but they can become verbs if merged to little $v$. Dispensing of roots with gender allows us to have a real lexical core that can fluctuate between syntactic categories without the need to store information multiple times.

Even if the storage of gender features on little $n$ is not entirely unproblematic, it seems the most plausible alternative to the three analyzed in this section and, therefore, it is the solution assumed in this dissertation. Nevertheless, it should be made clear that it is irrelevant whether we assume gender to be stored on little $n$ or on the root, as the realization of gender still ends up in the theme position in either case, which is what really matters for the proposal.[42]

---

[42]Locality conditions also seem to support the idea that gender features are housed on $n$, as they place gender in an intimate local structural relationship with the theme position —where the phonological features associated with gender are finally spelled out.

# FIVE

# PALLARESE 3RD PERSON CLITICS AND DEFINITE ARTICLES

## 5.1  Pallarese Catalan 3rd person clitics

In chapter 2 we saw the overall distribution of Pallarese pronominal clitics as well as definite articles. In this chapter, bearing in mind the previous proposals for gender inflection presented in chapter 3 and the assumptions made for the present analysis in chapter 4, I will focus on 3rd person accusative clitics not only because they exhibit special characteristics in comparison with 1st and 2nd person clitics but also because they differ substantially from the 3rd person clitics of other Catalan dialects.

Crosslinguistically, there is a split between 1st and 2nd person pronouns versus 3rd person pronouns. Benveniste (1966) regards the 3rd person as the 'non-person' because it does not participate as an active referent in the discourse, contrary to 1st and 2nd persons. This asymmetry is captured by the fact that in many languages the 3rd person, as it is less specified, needs overt gender marking because it does not bear person features (Harley & Ritter, 2002).

In many languages the form of 3rd person pronouns is taken from demonstratives, and this is in fact also the case in Pallarese Catalan, where definite articles and 3rd person accusative clitics share the exact same shape, derived from the Latin demonstrative *ille-illa-illud* 'that'. Their distribution, together with their phonetic variants, can be seen in table 5.1:

|            | Singular   | Plural       |
|------------|------------|--------------|
| **Feminine**  | [la]       | [ls], [les]  |
| **Masculine** | [l], [lo]  | [ls], [les]  |

Table 5.1: *Pallarese Catalan pronominal clitics and definite articles*

As shown above, 3rd person clitics are made up of a morph /l/ plus the realization of gender at its right edge. Although clitics are different from nouns —which they always refer to— there is a clear relationship between the morphosyntax of nouns and clitics. However, their singularity raises a question: is their morphological structure exactly the same as in nouns? Oltra-Massuet & Arregi (2005, p. 67) believe that it is:

> There are four types of nonverbs that we have not discussed so far: adverbs, determiners, strong pronouns, and nonfinite verbs. These words have a class marker, that is, a theme position whose realization is governed by the same principles as in nouns and adjectives. (...) Determiners and pronouns are of category D (see Abney 1987). Like any other functional category in Spanish, Th is adjoined to D at MS, as illustrated by the class markers in *múch-o* 'much', *ést-e* 'this (M)', *éll-a* 'she', *aquél-Ø* 'that (M)'.

Therefore, the only difference between nouns and clitics is their category-defining head, with D being the head for the latter. As already stated in previous chapters, all functional heads require a theme position, and thus in clitics it is the D head which projects this position, not little *n*, as illustrated in (1):

(1)

```
              D
             ╱‾‾╲
           D     Th
           |      |
           l    Ø/o/a
```
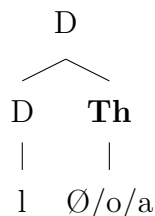
Clitics, however, are merely functional words and thus neither bear any lexical content nor carry gender features. Instead, they obtain them by agreement with the nouns they are referents for. Nonetheless, the theme position and all the principles that govern noun structures are maintained.

Given the structure in (1), it still needs to be determined what kind of semantic information the morph /l/ provides in the case of both articles and clitics. They

encode deixis, since they are referential links between linguistic expressions and their contextualization in the discourse (Eguren, 1999), but only anaphoric, not spatial as demonstratives. If articles and pronominal clitics have exactly the same form, though, they must have some common feature, and this feature is definiteness, as argued by Leonetti (1999) for Spanish or Martín (2012) for Catalan, among other authors. The structure of the clitic is thus the following:

(2)     Definiteness   +   gender (realized in the theme position)
        l            +   Ø/o/a

The feature composition of the clitic in (2) will be assumed throughout the rest of this dissertation. In sections 5.1.1 and 5.1.2, I present an analysis of both singular and plural clitics based on this idea together with the morphological structure posited in (1) in an attempt to answer the following question: *what* is inserted (and, crucially, *how* is it inserted) in the theme position?

## 5.1.1   Gender asymmetries: masculine vs. feminine

In Pallerese Catalan, singular masculine and feminine clitics show a different distributional pattern in postvocalic and postconsonantal contexts, as shown in table 5.2:

|  | **Singular** |
|---|---|
| **Feminine** | [la] |
| **Masculine** | [l], [lo] |

Table 5.2: *Singular clitics (postvocalic and postconsonantal)*

Whereas in the masculine the presence of a previous vocalic host alters the shape of the clitic, which does not bear an overt gender marker, as in (3-a) and (4-a), in the feminine the clitic surfaces always as [la], as in (5-a) and (6-a):

(3)   *Pron. accusative clitic (masc.)*

   a.   No  **[l]**       porto
      NEG ACC.3S.MASC bring.1S
      'I don't bring it'

   b.   **[lo]**       porto
      ACC.3S.MASC bring.1S
      'I bring it'

(4)   *Definite article (masc.)*

   a.   Agafa      **[l]**  pa!
      TAKE.IMP.2S DEF.M bread
      'Take the bread!'

   b.   Agafem   **[lo]**  pa
      TAKE.1PL DEF.M bread
      'We take the bread'

(5)  *Pron. accusative clitic (fem.)*

    a.  No  [**la**]  porto
        NEG ACC.3S.FEM bring.1S
        'I don't bring it'

    b.  [**la**]  porto
        ACC.3S.FEM bring.1S
        'I bring it'

(6)  *Definite article (fem.)*

    a.  Agafa  [**la**]  gerra!
        take.IMP.2S DEF.F vase
        'Take the vase!'

    b.  Agafem  [**la**]  gerra
        take.1PL DEF.F vase
        'We take the vase'

In absolute word-initial and postconsonantal contexts, when the word that follows the clitic is consonant-initial, the full form of the clitic surfaces. The Sonority Hierarchy and the specific phonotactic requirements of Catalan do not allow the syllabification of the clitic within a complex cluster in (3-b) and (4-b) for the masculine. That is, sequences like *[l p]*orto* 'I bring it' or *[l ʒ]*erra* 'the vase' are not permitted in Catalan, and thus an output without the inflectional ending is never allowed. The feminine form is also [la] in these contexts, (5-b) and (6-b).

By contrast, the output of both the masculine and the feminine is the non-vocalic form in prevocalic contexts, as in (7), (8), (9) and (10), where the clitic is always [l].

(7)  *Pronominal Feminine*

    a.  [**l**]'  agafo
        ACC.3S.FEM take.1S
        'I take it'

    b.  No  [**l**]'  agafo
        NEG ACC.3S.FEM take.1S
        'I don't take it'

(8)  *Pronominal Masculine*

    a.  [**l**]'  agafo
        ACC.3S.MASC take.1S
        'I take it'

    b.  No  [**l**]'  agafo
        NEG ACC.3S.MASC take.1S
        'I don't take it'

(9)  *Feminine definite article*

    a.  Agafem  [**l**]'  amanida
        take.1PL DEF.FEM salad
        'We take the salad'

    b.  Agafa  [**l**]'
        take.IMP.2S DEF.FEM
        amanida!
        salad
        'Take the salad!'

(10)  *Masculine definite article*

    a.  Agafem  [**l**]'  oli
        TAKE.1PL DEF.MASC oil
        'We take the oil'

    b.  Agafa  [**l**]'
        TAKE.IMP.2S DEF.MASC
        oli!
        oil
        'Take the oil!'

Pallarese clitics surface forms are therefore phonologically conditioned, exhibiting the following distribution:

|  | *Prevocalic* | *Postvocalic* | *Preconsonantal* |
|---|---|---|---|
| **Masculine** | [l] | [l] | [lo] |
| **Feminine** | [l] | [la] | [la] |

Table 5.3: *Phonological conditionings of Pallarese singular clitics*

I argue in this thesis that one single underlying form, consisting of the definiteness morph /l/ plus the floating phonological features corresponding to gender, can account for all the differences of table 5.3. Before moving on to the formal analysis, though, let me make clear again the lexical entries I assume for Pallarese Catalan gender inflection in (11), shown previously in chapter 4:[1]

(11)   a.   Underspecified vowel in the theme position:
Th ↔ V (*à la* Roca 2010)

b.   Association of gender values with phonological features:

| [+fem] | ↔ | [+low] | (= /a/) |
|---|---|---|---|
| [–fem] | ↔ | [+labial, –high] | (= /o/) |

Just as we saw in nouns, the theme position is spelled out as a [–cons] feature linked to a timing unit (i.e., an underspecified vowel). Gender, by contrast, is spelled-out as floating features, which will be linked or not to timing units only under certain conditions (as opposed to the vowel without place features that occupies the theme position). The idea of inflectional endings as floating features is not new in Romance languages, and, in fact, the OT account I will offer below has some similarities with del Gobbo (2001)'s analysis of Italian definite and indefinite articles sketched in section 4.2.

Del Gobbo argues for a constraint DEP-X ("every X slot of the output has an identical correspondent in the input") that prevents the insertion of floating vowels, since she considers that the gender marker *-o* is floating in masculine indefinite and definite articles —*un(o)* and *l(o)*— and only inserted when forced to do so by phonotactics.[2] Instead of DEP-X, I will make use of two DEP constraints to explain the behavior of clitics:

---

[1]As already stated in chapter 4, only the distinctive phonological features for the vowels *-a* and *-o* are provided in (11). Throughout this dissertation, [+labial] will be used as a shortcut for the features that correspond to /o/.

[2]Del Gobbo takes this idea from Tranel (1996), where the constraints PARSE-RT ("don't delete a root node") and PARSE-X ("don't delete an X-slot"), which go against the non-realization of floating features, account for the differences of a set of French word-final consonants. The Catalan data I am presenting here fall into this same domain since the parsing or non-parsing of phonological features depends on skeletal positions.

(12)     a.    DEPLINK: every association line in the output has a correspondent in
               the input (*à la* Morén 2001).
               This constraint goes against the insertion of association lines, and any
               floating feature in the input which is pronounced in the output (i.e.,
               linked to an X-slot) will violate it.

         b.    DEP-V: every vowel in the output has a correspondent in the input.
               DEP-V goes against vowel epenthesis, but it crucially differs from
               DEPLINK in that inserting a full vowel implies inserting a new struc-
               tural position, new association lines and new features. Therefore, DEPLINK
               is a subset of DEP-V, and candidates that only epenthesize association
               lines will harmonically bound candidates with full epenthesis.

Although the phonological features corresponding to gender are floating and DEPLINK
militates against their phonetic interpretation, other faithfulness constraints must
ensure that they are sometimes linked in the outputs, as *la* and *lo* are found in
preconsonantal contexts. The parsing constraints of (13) guarantee the linking of
phonological features:[3]

(13)     a.    PARSE[+LABIAL]: [+labial] must be parsed into the phonological struc-
               ture.
         b.    PARSE[+LOW]: [+low] must be parsed into the phonological structure.

Provided with these two kinds of conflicting constraints that advocate or go against
the interpretation of phonological features, we have two possible scenarios, depicted
in (14) and (15):

---

[3]I assume Stray Erasure (McCarthy, 1979) for unsyllabified segments that are not phonetically
realized.

(14)    If DepLink >> Parse[+labial]: no pronunciation

[+labial]


V

(15)    If Parse[+labial] >> DepLink: pronunciation

[+labial]
|
|
|
|
V

In (14), the phonological feature is not linked to the V-slot of the theme position if DepLink outranks Parse[+labial] (or [+low] in the feminine) and, therefore, it cannot be pronounced. By contrast, when the parsing constraint dominates DepLink, the output candidate with the overt inflectional ending is the winner, as in (15).

Finally, we need a phonotactic constraint $\sigma$Struc that guarantees syllable well-formedness in Pallarese Catalan:

(16)    $\sigma$Struc: comply with the syllable structure requirements of Catalan.[4]

With all these considerations in mind, we can now turn to the formal implementation of the proposal. I will analyze masculine singular clitics in 5.1.1.1 and feminine singular clitics in 5.1.1.2.

### 5.1.1.1    Masculine

Let us first look at the cases where the clitic surfaces as non-vocalic, as exemplified in (17) and analyzed by means of the tableau in (18):

---

[4]$\sigma$Struc must be understood as a cover constraint including SonSeq, which penalizes falling sonority onset clusters and rising sonority coda clusters, and *ConsonantalNucleus, militating against syllabic nuclei that are non-vocalic.

(17)  No  [**l**]       porto
NEG ACC.3S.MASC bring.1S
'I don't bring it'

(18)   *Masculine, vocalic host*

| [+labial]<br><br>/n  o  #  l  **V**  #  p  ɔ  r  t  o/ | σStr | Dep-Lk | Parse[+lab] |
|---|---|---|---|
| [+labial]<br><br>a. ☞  n  o   l     p  ɔ́  r  t  o | | | 1 |
| [+labial]<br>    &#124;<br>b.   n  o   l  o   p  ɔ́  r  t  o | | 1W | L |

In (18), σSTRUC is not active because the presence of the negation adverb *no*, with a vocalic nucleus, permits the syllabification of the clitic as a coda ([nol.pɔ́r.to]). As DEPLINK dominates PARSE[+LABIAL], candidate *a.* is the winner although the [+labial] feature is not phonetically interpreted.

A different clitic surfaces when it appears in preconsonantal contexts —the tableau in (20)— and there is no previous vocalic host to be attached to, as in (19):

(19)  **[lo]**      porto
      ACC.3S.MASC bring.1S
      'I bring it'

(20)  *Masculine, no host* (phonotactic problem)

| [+labial] /l **V** # p ɔ r t o/ | DEP-V | σStr | DEP-LK | PARSE[+LAB] |
|---|---|---|---|---|
| [+labial]<br>a. ☞  l  o    p  ɔ́  r  t  o |  |  | 1 |  |
| [+labial]<br>b.   l      p  ɔ́  r  t  o |  | 1W | L | 1 |
| [-back]  [+labial]<br>c.   e  l    p  ɔ́  r  t  o | 1W |  | 1 | 1 |

The contact between the clitic morph /l/ and the plosive /p/ of *porto* rules out candidate *b.*, which violates high-ranked σStruc. Although DEPLINK dominates PARSE[+LABIAL], some vowel needs to be inserted to avoid a phonotactic problem, and the parsing of [+labial] into the V-slot of the theme position is the more parsimonious option. That is, candidate *c.* violates DEPLINK anyway and, moreover, it violates DEP-V and PARSE[+LABIAL] as well. The creation of new material (features, association lines and X-slots) is a last resort option, whereas the use of pre-existing phonological objects is favored. In a way, it can be argued that the grammar 'recycles' what is already present in it —a strategy also used in syntax (Longa et al., 1998)— and this process could even be linked to Lexical Conservatism (Steriade, 1999) as the grammar uses an already existing morph for phonological purposes.[5]

In prevocalic contexts —the tableau in (22)— where there are no phonotactic

---

[5]The idea of Lexical Conservatism, though, is usually invoked when there is a mismatch between phonological form and morphosyntactic function. That is, to consider it a proper case of Lexical Conservatism the masculine singular clitic should use the feminine exponent /a/ to repair syllable structure.

problems, the dominance of Parse[+labial] over DepLink favors candidate *a*. The parsing of *-o* does not improve syllable structure and, therefore, the clitic remains as non-vocalic. This is exemplified in (21) and analyzed in the tableau in (22):

(21)   [**l**]'          agafo
       ACC.3S.MASC take.1S
       'I take it'

(22)   *Masculine, prevocalic*

| [+labial] /l **V** # a g a f o/ | Dep-V | σStr | Dep-Lk | Parse[+lab] |
|---|---|---|---|---|
| [+labial] a. ☞ l  a ɣ á f o | | | | 1 |
| [+labial] \| b.  l o  a ɣ á f o | | | 1W | L |

#### 5.1.1.2  Feminine

The feminine differs from the masculine in postvocalic contexts, as in (23):

(23)   No [**la**]          porto
       NEG ACC.3S.FEM bring.1S
       'I don't bring it'

In (24), Parse[+low] is ranked above DepLink, which forces the parsing of the feature [+low]. Candidate *b.* is ruled out because it does not bear the feminine exponent, whereas in the masculine the order was reversed (DepLink dominated Parse[+labial]) and the non-vocalic clitic was the winner.

(24)     *Feminine, vocalic host*

| [+low]<br>/n  o  #  l  **V**  #  p  ɔ  r  t  o/ | σStr | Parse[+low] | Dep-Lk |
|---|---|---|---|
| [+low]<br>     \|<br>a. ☞  n  o  l  a  p  ɔ́  r  t  o | | | 1 |
| [+low]<br>b.  n  o  l  p  ɔ́  r  t  o | | 1W | L |

The situation does not change in preconsonantal contexts, as seen in (25):

(25)     **[la]**         porto
     ACC.3S.FEM bring.1S
     'I bring it'

We saw in (20) that syllable structure forced the parsing of the [+labial] feature in these contexts, but in the feminine PARSE[+LOW] ensures that the non-vocalic form is never the winner, as shown in (24). Thus, in (26), candidate *b.* is ruled out due to σSTRUC, but it could never win in any scenario, as the [+low] feature needs to appear in the output even when syllabification poses no problems.

(26)     *Feminine, no host*

| [+low]<br>/l  **V**  #  p  ɔ  r  t  o/ | σStruc | Parse[+low] | Dep-Link |
|---|---|---|---|
| [+low]<br>   \|<br>a. ☞  l  a  p  ɔ́  r  t  o | | | 1 |
| [+low]<br>b.  l  p  ɔ́  r  t  o | 1W | 1 | L |

The clitic without the overt gender exponent appears only in prevocalic contexts:

(27)   [**l**]'        agafo
       ACC.3S.FEM take.1S
       'I take it'

The reason why we find this form only in prevocalic contexts is hiatus avoidance. As shown in tableau (29) below, we need the intervention of another constraint, defined in (28), to fully understand the whole patterning of the feminine singular clitic:

(28)     *V.V: two consecutive heteromorphemic vowels are not allowed

The constraint *V.V correctly generates the winning candidate *a.*, as PARSE[+LOW] is still higher-ranked than DEPLINK but candidate *b.* incurs in a violation of *V.V, and is thus ruled out.

(29)     *Feminine, prevocalic*

| [+low]<br><br>/l **V** # a g a f o/ | *V.V | σStr | Parse[+low] | Dep-Link |
|---|---|---|---|---|
| [+low]<br><br>a. ☞ l    a ɣ á f o | | | 1 | |
| [+low]<br>\|<br>b.  l a   a ɣ á f o | 1W | | L | |

This hiatus resolution strategy in word combinations is widespread in the phonology of Catalan. Vallverdú (2002) notes that, in Central Catalan word combinations, when two schwas are in contact (both in unstressed positions) as in (30-a), one of them is —almost always— not pronounced. Although there is no schwa in the vowel inventory of Pallarese due to a different pattern of vowel reduction in Northwestern Catalan, one of the vowels is also deleted (e.g., *port*[a] [a]*ntiga* > [pòrt**a**ntíɣa]). In cases where the vowel quality differs, the final unstressed *-e* or *-a* of the first word (always *-ə* in Central Catalan due to vowel reduction) can end up unpronounced as well before a non-high stressed vowel, as in (30-b). Though being always an optional

process, it is more frequent for the vowel to be deleted in function words such as quantifiers (as in (30-b)), clitics or determiners.

(30)    a.    *porta antiga* 'old door' [pɔ̀rtəntíɣə]


        b.    *quatre orques* 'four killer whales' [kwàtɾɔ́rkəs]

Words like *vehement* 'vehement' [bəəmén] (Central Catalan), with two tauto-morphemic consecutive schwas, show that *V.V operates only in heteromorphemic contexts.

## 5.1.2    Number asymmetries: singular vs. plural

Pallarese plural 3rd person pronominal accusative clitics and definite articles show no gender differences:

|  | *Singular* | *Plural* |
|---|---|---|
| *Masculine* | [l], [lo] | **[ls]**, **[les]** |
| *Feminine* | [la] | |

Table 5.4: *Pallarese clitics*

Although the form is the same for the masculine and feminine, their shape is phonologically conditioned, as exemplified in (31) and (32):[6]

(31)    *Pronominal accusative clitics*

        a.    No  **[ls]**          porto,  les        gerres      grogues
              NEG ACC.3PL.FEM bring.1S DEF.FEM.PL vase.FEM.PL yellow.FEM.PL
              'I don't bring them, the yellow vases'

        b.    No  **[ls]**          porto,  les        gots
              NEG ACC.3PL.MASC bring.1S DEF.MASC.PL glass.MASC.PL
              grocs
              yellow.MASC.PL
              'I don't bring them, the yellow glasses'

        c.    **[les]**          porto,  les        gerres      grogues
              ACC.3PL.FEM bring.1S DEF.FEM.PL vase.FEM.PL yellow.FEM.PL
              'I bring them, the yellow vases'

        d.    **[les]**          porto,  les        gots        grocs
              ACC.3PL.MASC bring.1S DEF.MASC.PL glass.MASC.PL yellow.MASC.PL
              'I bring them, the yellow glasses'

---

[6]I am using MASC and FEM in the glosses in order to emphasize whether the plural clitics have masculine or feminine antecedents. However, as noted, gender is not specified in their phonological shape.

(32)  *Definite articles*

    a.  Agafa    **[ls]**    gerres    grogues!
        take.IMP.2S DEF.FEM.PL vase.FEM.PL yellow.FEM.PL
        'Take the yellow vases!'

    b.  Agafa    **[ls]**    gots    grocs!
        take.IMP.2S DEF.MASC.PL glass.MASC.PL yellow.MASC.PL
        'Take the yellow glasses!'

    c.  Agafem **[les]**    gerres    grogues
        take.1PL DEF.FEM.PL vase.FEM.PL yellow.FEM.PL
        'We take the yellow vases'

    d.  Agafem **[les]**    gots    grocs
        take.1PL DEF.MASC.PL glass.MASC.PL yellow.MASC.PL
        'We take the yellow glasses'

After the plural exponent /z/ is joined to the definiteness morph /l/, the clitic surfaces as the masculine singular under the same phonological conditionings.[7] That is, in postvocalic contexts, when the vowel of the preceding word can be the nucleus of a syllable, it surfaces as non-vocalic, as in (31-a) and (31-b) (pronominal clitic) or (32-a) and (32-b) (definite article). In preconsonantal contexts without any previous vocalic host, it appears with the full form [les] ((31-c) and (31-d) for pronouns and (32-c) and (32-d) for definite articles), as the non-vocalic form would create a phonotactic problem. Therefore, instead of the alternation [l]/[lo] of the masculine singular, we find the variants [ls]/[les], with a different vowel. It is impossible to understand the behavior of the plural clitics, however, without addressing the following questions:

1. Why are there no gender differences in the plural?

2. What is the nature of this vowel -*e* and why is it different from the vowels found in the singular clitics?

  The absence of gender features and, as a consequence, the absence of their corresponding phonological exponents is what determines the use of a different vowel in the plural. However, before understanding the mechanisms that select -*e* over other vowels and its location (i.e., in the theme position), we first need to account for the asymmetries found in number with the non-realization of gender in the plural.

  Crosslinguistically, there is a dependency of gender on number. Of particular interest are Greenberg's universals 37 and 45:

(33)    a.    "A language never has more gender categories in nonsingular numbers than in the singular." (Greenberg 1963, p. 95).

---

[7]For the reasons explained in 3.2 the plural morph must be underlyingly /z/ and not */s/.

b.   "If there are any gender distinctions in the plural of the pronoun, there are some gender distinctions in the singular also." (Greenberg 1963, p. 96)

Universal 45 is corroborated by Corbett (2000), who shows that if a language has gender distinctions in the plural, it will have them also in the singular. The opposite notion (i.e., the implication that a language having gender distinctions in the singular must have them in the plural) does not hold because in Pallarese masculine and feminine are specified in singular clitics but not in plural ones. But why is this the case and not the other way around? Greenberg's universals actually rely on markedness, because —without getting into the debate about the nature of markedness and its different dimensions— it is obvious that plural is more marked than singular (in complexity terms), since the former includes the latter (e.g., the notion of a single cat is implied in the plural 'cats').[8] In the literature, this kind of markedness has been labelled *Marked Features Restrict Subdistinctions* (Nevins, 2006). That is, in marked environments, the different feature specifications of certain categories can be neutralized. This is exactly what happens in Pallarese, as gender differences are maintained in the singular but not in the plural.

Following the idea first introduced by Bonet (1991), I understand this 'restriction on subdistinctions' as feature Impoverishment. Bonet proposes that, postsyntactically, in the morphological component, certain features are deleted and thus cannot be phonologically spelled out (in a late insertion model, the exponents are inserted after morphological operations such as Impoverishment, Fusion or Fission). While these features are active in the syntactic derivation, their postsyntactic deletion restricts the number of phonologically overt morphological contrasts. This inevitably leads to syncretism, but of course this cannot be an unconstrained outcome of the grammar as oppositions are crucial for linguistic interpretation. If Impoverishment does not apply freely, what triggers it? Many authors (Bonet 1991, Noyer 1992 and Arregi & Nevins 2007, among others) have shown that markedness, as already noted, is responsible for this. Nevins (2006)'s *Marked Features Restrict Subdistinctions* can thus be reformulated as *Marked Features Trigger Impoverishment*, and in Pallarese Catalan the particular marked feature is [+plural]. In the singular, the unmarked value, maintaining a gender difference does not pose any problem, but, when it comes to the plural, neutralization takes place.

---

[8]Markedness is a multidimensional concept and need not pressupose the idea of complexity as in the singular-plural distinction. In fact, masculine is unmarked with respect to feminine and, conceptually, it is not less complex. Markedness refers in this case to default patterns which do not necessarily need to be more complex.

Nevertheless, plural does not always trigger gender neutralization. In fact, Impoverishment is only restricted to definite articles and 3rd person accusative clitics. In other determiners, adjectives, nouns and strong pronouns, the categories which can be inflected for gender and number, the distinction is maintained in both the singular and plural, as shown in (34) and (35):

(34)  *Indefinite article*

    a.  [una]    nevada
        a.FEM.S snowfall.FEM.S
        'A snowfall'

    b.  [un]    home
        a.M.S man.M.S
        'A man'

    c.  [unes]    cases
        a.FEM.PL house.FEM.PL
        'Certain houses'

    d.  [uns]    cosins
        a.M.PL cousin.M.PL
        'Certain cousins'

(35)  *Demonstratives*

    a.  [aɣéta]    setmana
        this.FEM.SG week.FEM.S
        'This week'

    b.  [aɣét]    ferrat
        this.M.SG bucket.M.S
        'This bucket'

    c.  [aɣétes]    portes
        this.FEM.PL door.FEM.PL
        'These doors'

    d.  [aɣéts]    homens
        this.M.PL cousin.M.PL
        'These men'

The examples above make it clear that not all the categories that bear [+plural] undergo Impoverishment and, in fact, it only applies to a very restricted set of functional words. As mentioned above, the defining feature of these functional elements, i.e., definite articles and 3rd person accusative clitics, is definiteness, realized as /l/. Given that only the conjunction of [+definiteness] and [+plural] triggers the deletion of gender features, we can postulate an Impoverishment rule as in (36):[9]

(36)     [±fem]⟶ Ø/[+def]___[+pl]

After this rule has applied, then, only [+plural] and [+definiteness] are part of the featural composition of the clitic. As a consequence, when Vocabulary Insertion takes place, the phonological exponents that match gender features cannot be inserted due to the lack of such features and, therefore, [+low] (feminine) and [+labial]

---

[9]A "conservative" view on Impoverishment is adopted here, as rules like (36) *delete* syntactic features in the morphological component. Some authors, like Trommer (2002), argue instead that Impoverishment should be viewed as a morphological constraint that bans the realization of certain syntactic configurations. This is also the idea defended by Bermúdez-Otero (2012, p. 53), who notes: "For example, a process of impoverishment enforcing a systematic syncretism (see e.g. Albright and Fuß this volume) can no longer be literally formulated as a *rule deleting syntactic features from syntactic output representations, but must rather be stated as a morphological constraint forbidding the use of exponents specified for a certain syntactic feature in the context of certain other features.* This is precisely the line taken in Wunderlich and Fabri's (1995) Minimalist Morphology and Trommer's (2001: e.g. 113) Distributed Optimality. [emphasis mine]"

(masculine) can never surface in the plural clitic. However, these features must be
realized in the theme position, and this position has not been impoverished. In
other words, the deletion of gender features does not imply the loss of the theme
position where they are realized, as the postsyntactic well-formedness requirement
that creates a theme position for all functional heads is still completely dissoci-
ated from Impoverishment (a different postsyntactic operation).[10] This assumption
has implications for lexical insertion (for clarification purposes, I repeat once again
the Vocabulary Items assumed for gender and the theme position stated in (11),
because, although (37-b) is inactive, the theme position is anyway filled with an
underspecified vowel —(37-a):

(37)    a.    Underspecified vowel in the theme position:
              Th ↔ V (*à la* Roca 2010)
        b.    Association of gender values with phonological features:

| | | | |
|---|---|---|---|
| [+fem] | ↔ | [+low] | (= /a/) |
| [–fem] | ↔ | [+labial, –high] | (= /o/) |

The morphological structural position has not been deleted and a [–cons] feature
associated with a timing unit is inserted there. As there is no context for insertion,
-*o* ([+labial]) and -*a* ([+low]) cannot end up in that position, but it still needs to be
determined why -*e* is the final exponent and why it appears only in phonotactically
problematic sequences.

Pallarese has a vowel inventory consisting of seven vowels in stressed positions
and five vowels in unstressed positions. In such systems, it is quite common to have
-*e* as an epenthetic vowel, and this is actually what we see in Pallarese, a dialect
that establishes a mid front vowel as the default.[11] This has to do with perceptual
prominence and positional markedness (Steriade, 1997) and explains why schwa,
the least perceptually salient vowel, is usually chosen in epenthetic contexts when
available, as it represents the unmarked option. But why does Pallarese, which lacks
schwa in its inventory, choose -*e* as the default epenthetic vowel? I follow Lloret &
Jiménez (2008)'s analysis of epenthetic vowel quality in Catalan in assuming that
there is a tension between strong positions (e.g., syllable nuclei, stressed syllables,
radicals, etc.), which demand open (i.e., more perceptually salient) vowels, and

---

[10]Impoverishment and the creation of the theme position are both postsyntactic and pre Vocab-
ulary Insertion operations. There is no need to order them in a specific way since conflict between
them is impossible. When Vocabulary Insertion takes place, there will be no gender features due to
Impoverishment and, as a consequence, no phonological realization of either masculine or feminine.
The theme position, however, is created anyway.

[11]In certain environments, however, the epenthetic vowel of Pallarese —and Northwestern Cata-
lan in general— can switch to -*a* (see Jiménez 2002, section 4.1.1.1, for more details).

weak positions (syllable margins, posttonic syllables, affixes, etc.), that prefer less perceptible vowels. The harmony hierarchy shown in (38) captures the adequacy of vowels to be syllable nuclei:

(38)     N/a  ≻  N/ɛ,ɔ  ≻  N/e,o  ≻  N/i,u  ≻  N/ə

We can see in (38) that /a/ is the most preferred vowel to become the nucleus of a syllable and /ə/ the least. However, this situation changes in syllable margins and other weak positions. Lloret & Jiménez (2008) argue that, in Western Catalan (the macrodialect Pallarese belongs to), the situation of the vowel in the Prosodic Word plays a role in determining its quality. That is, in Valencian and Northwestern Catalan, word-finally the epenthetic vowel is /e/, while word-initially it is /a/ due to a higher psycholinguistic prominence of this position. It is clear, then, that there is a tension between ideal syllable nuclei and weak positions, thus creating the asymmetry between word-initial and word-final epenthesis. Clitics are weak elements (they do not bear stress, for example, and do not have lexical content), so it would be logical to use /i/-/u/ as epenthetic elements according to the hierarchy in (38) (as schwa is absent in the phonology of Pallarese). However, these vowels are worse syllable nuclei than /e/-/o/ and are thus ruled out. Assuming that [+labial] is a marked feature that rules out /o/ as a possible option (as De Lacy 2002 points out, all candidates that contain the feature [+labial] are harmonically bounded by their non-labial counterparts), we can then establish /e/ as the default epenthetic vowel of Pallarese.

At first sight, then, it seems plausible to assume that the vowel *-e* of the plural clitic is epenthetic in nature. Nonetheless, the epenthetic patterns of Catalan that were shown in 4.1.2, which are sketched below again, seem to contradict this idea:

- Epenthesis applies word-initially

- Epenthesis does not apply word-finally because it interferes with the theme position of the noun system, creating mismatches between morphological and phonological structures

The theme position of the clitic thus disallows the insertion of an epenthetic vowel, ruling out the possibility of 'pure' epenthesis. Still, the contact between the definiteness morph /l/ and the plural morph /z/ creates a phonotactic problem. As noted above, there are no phonological features linked to gender that can be used to improve syllable structure and it cannot be argued either that the vowel of the clitic is lexicalized because there is a [Ø]∼[e] alternation that depends on the phonological context and is truly systematic. Recall, however, that the theme position with a

[–cons] feature linked to a timing unit is already spelled out and any inserted feature will end up there. My proposal is that the features provided by the phonological component are epenthetic in nature but dependent on the theme position, something that resembles Harris (1999)'s analysis of the Spanish nominal system (see section 3.1.2). According to him, in nouns like *hombre* 'man', the features of the vowel are epenthesized but finally interpreted as a class marker because they occupy a thematic position. I am not assuming this for Catalan, as in my analysis the vowel ending of *centre* 'center', parallel to that of *hombre* in Spanish, is lexical in nature, not epenthetic, but for the plural clitics the situation is the same, although the morphological structure differs from Harris's. The fact that the clitic forms are phonologically conditioned is crucial for adopting this view; that is, we have two different forms for the clitic, *ls/les*, whereas in *hombre* there is no [Ø]∼[e] alternation. It seems more plausible to assume, then, that -*e* is lexicalized in the latter case but not in the former. Tableaux (40) and (42), below, show the alternation [ls]/[les] in the plural clitics, with an input, in both cases, that does not contain any floating features, as opposed to the masculine and feminine singular clitics seen in section 5.1.1. The fact that gender has been impoverished in the plural blocks the realization of the phonological features that correspond to it. Nevertheless, the theme position has been spelled out anyway by a [–cons] feature that occupies a timing unit, but no place features are linked to the root node to determine the quality of the vowel. If the clitic is in phrase-initial position, as in (39), the contact between the clitic and the following verb forces the use of -*e*, as shown in (40). The choice of epenthetic vowel quality is captured by the constraints DEP[+HIGH] (against the insertion of /i/) and DEP[-BACK] (against the insertion of /e/), which actually reflect the hierarchy in (38). For explanatory reasons I am using these constraints that go against the insertion of certain features, but the overall picture is much more complicated, as I would need to integrate indexes for specific positions (as noted above, the difference between weak and strong structural positions is important, and this is why we do not find /i/ as the default epenthetic element). I am simply ordering DEP[+HIGH] above DEP[-BACK] to guarantee that /e/ is the winning candidate, but the considerations just noted should be borne in mind in order to fully understand the choice of vowel quality:

(39)  [les]              porto
      ACC.3PL.FEM/MASC bring.1S
      'I bring them'

(40)  *Plural, no host (phonotactic problem)*

| /l **V** + z # p ɔ r t o/ | Dep-V | σStr | Dep[+hi] | Dep[-bk] | Dep-Lk |
|---|---|---|---|---|---|
| a. ☞  [-back]  l  e  s    p  ɔ́  r  t  o |  |  |  | 1 | 1 |
| b.  [+high]  l  i  s    p  ɔ́  r  t  o |  |  | 1W | L | 1 |
| c.  l  s    p  ɔ́  r  t  o |  | 1W |  | L | L |
| d.  [-back]  e  l  s    p  ɔ́  r  t  o | 1W |  |  | 1 | 1 |

When there is no previous vocalic host, as in (40), the situation parallels that of the masculine singular clitic. The vocalic quality is different because, as mentioned above, morphology cannot provide the phonological features that correspond to gender values after gender has been impoverished. Instead, [–back] is epenthesized and linked to the V-slot of the theme position because σSTRUC forces the use of a vowel to avoid the onset cluster *lsp* of candidate b. Again, as happened with the preference for *lo* over *el* in the masculine singular clitic, candidate c. is ruled out because epenthesizing phonological features, association lines and structural positions is more costly than inserting only a phonological feature (with its corresponding association line). That is, in candidate c. the whole vowel needs to be epenthesized, the timing unit included, whereas in the winner the timing unit has been already filled with a V-slot by the morphological component and only the features that will end up there need to be epenthesized.

The situation changes when there is a vocalic host, as in (42), paralleling, again, the behavior of the masculine singular clitic. The vowel of the adverb *no* allows the syllabification of the plural clitic without the insertion of a vowel, as there is no phonological motivation for it:

(41) No [ls] porto
NEG ACC.3PL.FEM/MASC bring.1S
'I don't bring them'

(42) *Plural, vocalic host*

| /n o # l V + z # p ɔ r t o/ | Dep-V | σStr | Dep[-hi] | Dep[-bk] | Dep-Lk |
|---|---|---|---|---|---|
| a. ☞ n o l s p ɔ́ r t o | | | | | |
| [-back] <br> b. n o l e s p ɔ́ r t o | | | | 1W | 1 |
| [+high] <br> c. n o l i s p ɔ́ r t o | | | 1W | | 1 |
| [-back] <br> d. n o e l s p ɔ́ r t o | 1W | | | 1 | 1 |

Thus, in presence of a vocalic host, candidate b. is ruled out because it has epenthesized [–back] —the feature that corresponds to the vowel *-e*— unnecessarily (both the definiteness and the plural morphs can be attached to the adverb *no*, [nols]), which likewise implies the violation of DepLink, as the association lines that link it to the V-slot have to be epenthesized as well. Candidate c. is ruled out by the same reasoning, and also because the epenthetic place features of *-i* are dispreferred over *-e* (candidate d. is even less optimal because it epenthesizes a whole vowel, not only place features). By contrast, candidate a. does not incur in any violation because it respects the syllabic requirements of Pallarese and does not epenthesize phonological features or association lines. Without a place node, the [–cons] feature cannot be interpreted by the phonetic component and, thus, it is finally not pronounced. Obviously, Parse[+labial] and Parse[+low] do not play any role in this tableau after Impoverishment of gender features.

The data presented so far clearly indicate that epenthesis in Pallarese is conditioned by morphological structure, namely the theme position where gender is realized. The proposal has dealt thus far with 3rd person accusative pronominal clitics and definite articles, but it will be shown in the next section, 5.1.3, that the same analysis can be extended to 3rd person dative clitics.

### 5.1.3   Case asymmetries: dative vs. accusative

Pallarese dative (plural) clitics resemble 3rd person accusative clitics, as shown in table 5.5 below:[12]

| *Singular* | *Plural* |
|:---:|:---:|
| [li] | |
| [li] | [ls], [les] |

Table 5.5: *Pallarese dative clitics*

There are gender differences neither in the singular nor in the plural, although in the singular there is one single form [li] in all contexts (pre- or postvocalically), while in the plural there are two different phonologically conditioned forms, [ls] and [les]:

(43)   *Singular*

    a.   **[li]**   agafa   la   mà
       DAT.3S take.3S the.FEM.S hand.FEM.S
       'S/he takes her/his hand'

    b.   No **[li]**   agafa   la   mà
       NEG DAT.3S take.3S the.FEM.S hand.FEM.S
       'S/he doesn't take her/his hand'

(44)   *Plural*

    a.   **[les]**   toques   la   mà
       DAT.3PL touch.2S the.FEM.S hand.FEM.S
       'You touch their hand'

    b.   No **[ls]**   toques   la   mà
       NEG DAT.3PL touch.2S the.FEM.S hand.FEM.S
       'You don't touch their hand'

It is clear in the examples above that the plural dative clitic has exactly the same shape as the accusative (and the definite article). Only sporadically does a marker /i/, always present in the singular, appear after the number morph:

(45)   I'   **[lz i]**   donaven   pienso   natural
       and DAT.3PL give.IMPF.3PL feed.M.S natural.M.S
       'And they gave them natural feed'

Crucially, this morph is (almost) always absent in isolation but sometimes appears

---

[12]The shape of the singular dative clitic is *a priori* not the same as in the accusative singular, but in this section it will be proven that before Impoverishment takes place, their underlying syntactic structure is equivalent, except for some deictic information attached to the dative forms.

in combinations (in isolation, the form *les hi* seems to appear due to the influence of Central Catalan, as the most conservative speakers —those who show more 'genuine' dialectal features— do not use it). I will talk about this difference later on in section 5.1.3.1, as the nature of /i/ needs to be addressed in order to account for the distribution of the 3rd person dative in singular and plural forms, either in isolation or in combinations.

As already mentioned, I am assuming Martín (2012)'s decomposition of 3rd person clitics, where the /l/ morph is considered to encode definiteness. In his proposal, accusatives are composed of definiteness plus gender features (which are realized in the theme position under the present analysis), whereas datives are, in turn, "complex grammatical entities that contain accusative clitics as one of their several constituent parts" (Martín, 2012, p. 155). But why should datives be considered 'complex entities'? Although a full answer to the question is outside the scope of this dissertation, the phrase-level and clitic examples (46)–(49) from Martín (2012, p. 156-157) justify this choice:

(46)  *Accusative (phrase level)*

    a.  Veig  [la  noia]$_{acc}$
        see.1S the.FEM.S girl.FEM.S
        'I see the girl'

    b.  Je vois  [la  fille]$_{acc}$
        I  see.1S the.textsc.f.s girl.FEM.S
        'I see the girl'

(47)  *Accusative clitic*

    a.  La  veig
        ACC.3S see.1S
        'I see her'

    b.  Je la  vois
        I  ACC.3S see.1S
        'I see her'

(48)  *Dative (phrase level)*

    a.  Li  dono els  llibres  [a [la  noia]$_{acc}$]$_{dat}$
        DAT.3S give.1S theMASC.PL books.MASC.PL to the.FEM.S girl.FEM.S
        'I give the books to the girl'

    b.  Je donne les  livres  [à [la  fille$_{acc}$]$_{dat}$
        I  give.1S the.MASC.PL books.MASC.PL to the.FEM.S girl.FEM.S
        'I give the books to the girl'

(49)  *Dative clitic*

  a. Li   dono  els    llibres
     DAT.3S give.1S the.MASC.PL books.MASC.PL
     'I give her/him the books'

  b. Je lui  donne les    livres
     I DAT.3S give.1S the.MASC.PL books.MASC.PL
     'I give her/him the books'

What these examples show is that the difference between the dative —(48)— and the accusative —(46)— at the phrase level is expressed by the use of the preposition *a/à*.[13] This must be incorporated into the dative clitic, but the question is how. In any case, if we are assuming that the dative is complex in nature and already contains the accusative, it must look as in (50), according to Martín:

(50)  DATIVE = ACCUSATIVE + $x$

So what is the other consituent labeled $x$? Whatever replaces the preposition *a* must incorporate somehow its directional/spatial meaning. For this reason, Martín (2012) assumes that the morpheme that complements the accusative expresses deixis, an idea taken from Kayne (2008), who connects the morpheme present in existential constructions and datives alike with a deictic marker. Deixis is thus expressed by Catalan *hi*, Paduan *ghe*, Italian *ci* or French *y*, in both existential —(51)— or dative constructions —(52):

(51) a. Ghe ze un   libro
    DX be3S a.MASC.S book.MASC.S
    'There is a book'

  b. Hi ha  un   llibre
    DX have.3S a.MASC.S book.MASC.S
    'There is a book'

(52) a. Ghe  dago il   libro
    DX/DAT give.1S the.MASC.S book.MASC.S
    'I give them/her/him the book'

  b. Els  hi  dono els    llibres
    ACC.3PL DX/DAT give.1S the.MASC.PL book.MASC.PL
    'I give them the books'

 The general idea, then, is that the information in both datives and locative or existential constructions provided by the morph /i/ is the same, namely deixis. In Martín (2012, p. 220), this is directly incorporated into the syntactic structure of

---

[13]Martín (2012, p. 158) notes that *a* is used in Catalan and Spanish in all dative phrases but only in certain accusative structures (in Differential Object Marking contexts for human accusative objects), suggesting that "accusative is a subset of dative".

the dative by a Deixis Phrase, as in (53):

(53)

```
        DP₂
       ⟋ ⟍
      l    DxP
          ⟋ ⟍
        hi    DP₁
             ⟋ ⟍
            l   NP
```
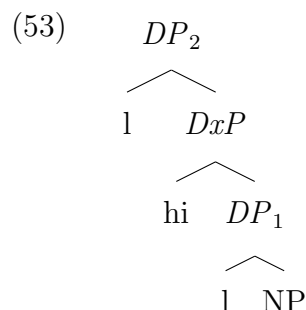
What I mean to emphasize choosing Martín's syntactic decomposition of the dative clitic is that *hi*, the head of the DxP, is encoded in the dative and is crucial for its interpretation. If the definiteness morph -*l* occupies the head of the lower $DP_1$, the clitic is interpreted as an accusative structure (with gender being inherited from the referent it stands for), but it is interpreted as a dative in $DP_2$ thanks to the deixis phrase. This structure is in principle applicable to all Romance clitics, but both the deictic and the definiteness morphs can be silent, as in Spanish (*le, les*) or Paduan *ghe* —originally a locative.[14]

Deixis must thus be encoded in the dative, whose final decomposition is shown in (54):

(54)     DATIVE  =  ACCUSATIVE  +  *deixis*

Given this template, we can now provide the final morphosyntactic structure for the clitic with every morph occupying a specific structural position (including the theme position):

|  | *Def.* | | *Th* | | *Number* | | *Deixis* | | |
|---|---|---|---|---|---|---|---|---|---|
| (55) *Singular*: | l | + | Ø | | | + | i | = | [li] |
| *Plural*: | l | + | e/Ø | + | z | + | (i) | = | [les]/[ls] |

What (55) shows is that another good argument for treating /i/ as a deictic marker comes from its morphological location. Number is usually the most peripheral feature in Catalan, but the morph [i] is joined after number in plural clitics, as in [lezi] (or [əlzi] in Central Catalan). If it were a 'regular' component of the dative it would appear before number, but its special status as an independent deictic marker allows its insertion after it. Worth mentioning as well is the combination of the dative plural *els* and the genitive *en* in Central Catalan, [əlzəni]. This sequence is a good

---

[14]Martin & Hinzen (2014, p. 110) also argue that the dative "can be lexicalized by any part of the complex dative phrase" in (53), all other parts being silent.

argument for the treatment of /i/ as an independent morpheme in the dative, because the genitive gets inserted between the proper dative morph and the deictic. The situation in Pallarese, though, is different, as will be shown later in this chapter.

Leaving aside this fact, two relevant questions need to be answered to account for the general patterning of dative clitics:

1. Where are gender features realized, if they are realized at all?

2. Does -*e* have the same status in the plural dative as in the accusative?

The segmentation in (55) clearly indicates that the theme position does not have any realization in the singular, while in the plural there is an alternation $e/\emptyset$ parallel to that of the accusative plural (actually, the exponence of accusatives and datives in the plural is the same, even when their syntactic configuration is different). If the accusative, as we saw in (53), is contained in the dative, why is gender not expressed in it? There is a certain amount of consensus in the literature that Gender and Person are in complementary distribution in the languages of the world (Greenberg, 1963; Harley & Ritter, 2002), and this is what encourages Martín (2012) and Hinzen & Sheehan (2014) to argue that gender features are interpreted somehow as person features due to the effect of the deictic features of the dative. While it is true that dative clitics are intuitively more connected to participants in the discourse than accusative clitics (which can refer to objects as well), it is not clear to me why they must then "enter the realm of personal clitics" (Hinzen & Sheehan, 2014, p. 68). Again, the 3rd person is the 'non-person', so treating 3rd person datives as personal due to the blocking effect of deixis in the interpretation of gender is not particularly convincing. Instead, my proposal captures the absence of gender features not only in dative vs. accusative but also in plural vs. singular (in the pronominal accusative clitics and definite articles), two apparently independent phenomena.

It was shown in (36) that markedness triggered Impoverishment and gender was not realized in the accusative plural, plural being a marked category. Going back to markedness (understanding markedness in complexity terms), we can unify the puzzle of number and case asymmetries regarding gender if we extend the Impoverishment rule of (36) to the dative clitic. If the hypothesis about the structural composition of the dative clitic is correct and the accusative is part of it, it follows that the dative is also more complex, as it is a superset of the accusative. What feature, though, triggers Impoverishment? The deictic features that I am assuming to be part of the dative allow us to postulate a more fine-grained Impoverishment rule in (56), which captures the behavior of the plural dative clitic (as well as the

accusative):[15]

(56)     $[\pm\text{fem}] \longrightarrow \emptyset/[+\text{def}]\underline{\quad} \left\{ \begin{array}{c} [+\text{pl}] \\ [+\mathbf{dx}] \end{array} \right\}$

In (56) it is stated that gender is deleted when there is co-occurence of definiteness and plural or deictic features (or both). This is precisely what happens in Pallarese, where we find the following distribution of 3rd person plural and dative clitics:

(57)     [li]      =   dative singular, masculine/feminine

         [ls],[les]  =   dative/accusative plural, masculine/feminine

As already noted, it is a fact that, crosslinguistically, gender differences are neutralized in the plural. In the plural dative that occupy us here, besides gender not being realized, the deictic morpheme /i/ that appears in the singular is also absent.[16] Why number should trigger all these changes at the morphological level is not only a truly intriguing question but also very difficult to answer, as the structural motivations for it remain opaque. Nonetheless, since it lies outside of the scope of this thesis to explore this issue, I will only speculate about the behavior of the deictic marker /i/ in isolation and in combinations, which is sketched below:

- In the singular it always appears

- In the plural it does not appear in isolation

- In combinations it always appears (in both singular and plural)

To account for these differences in the manifestation of deixis, it could be argued that, in the singular clitic, the deictic morpheme serves a phonological purpose as well, along the lines of the gender exponents that were used for syllabification in the masculine accusative clitic and definite article. Let me elaborate. After the Impoverishment rule in (56) has applied, the dative is left without gender features. Without the deictic marker, we would only have a definiteness morph /l/ and an empty theme position. Thus, we could imagine a dative clitic with an alternation *l/*le, where a default vowel —whose features had been provided by the phonological component— would occupy the theme position. This scenario, depicted in (58),

---

[15]Even if the deictic is (almost) never realized in the plural clitic, it is still part of its syntactic configuration and therefore acts as a target for Impoverishment.

[16]Bonet (p.c.) points out that there could be Impoverishment of the deixis feature in plural contexts as well. It is not clear to me, then, why it remains in combinations, although it might be used there to signal their special status (the fusion of accusatives and datives in one single form), as I will explain later in (85).

would parallel the situation of the accusative plural clitic with the phonologically conditioned forms [ls]/[les]:

(58)  a.  *no  **[l]**         porto      (la        pilota)
          NEG DAT.3S.MASC bring.1S the.FEM.S ball.FEM.S
          'I don't bring her/him the ball'

      b.  *[**le**]        porto      (la        pilota)
          DAT.3S.MASC bring.1S the.FEM.S ball.FEM.S
          'I bring her/him (the ball)'

If this were the case, the difference between the accusative and the dative would depend only on the realization of gender, since the repair strategy used in the accusative masculine singular uses the inflectional morph *-o*, as illustrated below:

(59)  a.  no  **[l]**       porto
          NEG ACC.3S.MASC bring.1S
          'I don't bring it'

      b.  [**lo**]       porto
          ACC.3S.MASC bring.1S
          'I bring it'

A form *\*l/\*le* would resemble the Spanish dative *le* (no phonological conditioning in Spanish, though), where there is no overt trace of a deictic (according to Martín 2012's proposal, it is silent). The real situation, however, is different in the Pallarese singular dative, with the deictic morph /i/ appearing under all circumstances:

(60)  a.  [**li**]        porto      (la        pilota)
          DAT.3S.MASC bring.1S the.FEM.S ball.F.S
          'I bring her/him (the ball)'

      b.  no  [**li**]       porto      (la        pilota)
          NEG DAT.3S.MASC bring.1S DEF.F.S ball.DEF.F.S
          'I don't bring her/him the ball'

      c.  [**li**]      agafa  la       mà
          DAT.3S take.3S the.FEM.S hand.FEM.S
          'S/he takes her/his hand'

In (60) we see that the clitic always surfaces as *li* in preconsonantal, prevocalic or postvocalic environments. As there is one single variant of the dative singular, the phonological hypothesis must be discarded. We are obliged to assume, instead, that the deictic is always incorporated into the phonological exponence of the dative, not specifically for syllabification purposes as in the case of the singular accusative 3rd person clitic *l/lo*.

As noted previously, it is not clear why this deictic element is absent in the plural

(in isolation), but it might be no coincidence that the use of /i/ would not solve illicit phonotactic sequences. Deixis appears after number, so a sentence like (61) would be unpronounceable with or without /i/:

(61)  *[lz]/*[lzi] dono   açò
      DAT.3PL give.1S this.NEUT
      'I give them this'

The only solution is to insert a vowel either word-initially, the strategy employed by Central Catalan ([ə]ls), or between the *l*-morph and the plural, as in Pallarese ([les]). In this case, which is exactly what happens in the accusative plural, once we have epenthesized the features that will end up occupying the theme position, the deictic marker cannot improve syllable structure. It could in a way, as a CV.CV sequence ([lezi]) is preferable in markedness terms, but a CVC syllable ([les]) is totally acceptable and pervasive in the phonology of Catalan. For this reason, we could argue that the deictic morph is not necessary in the plural for phonological purposes, although we have already ruled out the phonological solution in the singular dative *li* and thus it does not seem plausible either in the plural (we still need to account, however, for the asymmetry between singular and plural).

Without attempting to explain the motivations for the appearance of the morph /i/, what is clear is that it is not present in Pallarese plural dative clitics. As a matter of fact, as we already noted, the 3rd person dative plural shares the exact same shape as the 3rd person accusative plural and definite article, [les]/[ls]. The examples in (62) and (64) and their corresponding tableaux in (63) and (65) illustrate again what was shown for the plural in the accusative in 5.1.2:

(62)    No   [ls]        porto    la
        NEG  DAT.3PL  bring.1S  DEF.FEM.S
        pilota
        ball.FEM.S
        'I don't bring them the ball'

(63)    *Plural, vocalic host*

| /n  o  +  l  **V**  z  +  p  ɔ  r  t  o/ | σStr | Dep[-bk] | Dep-Lk |
|---|---|---|---|
| a. ☞  n  o     l  s     p  ɔ́  r  t  o | | | |
| [-back]<br>\|<br>b.  n  o     l  e  s     p  ɔ́  r  t  o | | 1W | 1 |

(64)    [les]       porto    la        pilota
        DAT.3PL  bring.1S  DEF.F.S  ball.F.S
        'I bring them the ball'

(65)    *Plural, no host (phonotactic problem)*

| /l  **V**  z  +  p  Ø  r  t  o/ | Dep-V | σStr | Dep[-bk] | Dep-Lk |
|---|---|---|---|---|
| [-back]<br>\|<br>a. ☞  l  e  s     p  ɔ́  r  t  o | | | 1 | 1 |
| b.  l  s     p  ɔ́  r  t  o | | 1W | L | |
| [-back]<br>\|<br>c.  e  l  s     p  ɔ́  r  t  o | 1W | | 1 | 1 |

The Gender Impoverishment rule of (56), which operates in the presence of both

[+plural] and [+deixis] features, renders the dative clitic with no gender features, and thus no floating phonological exponents associated with it will be at hand as a way of repairing syllable structure. Again, though, Impoverishment does not affect the theme position, which remains unaltered with its [–cons] feature associated with an X-slot. For this reason, the vowel is not pronounced in the winning candidate in (63) (the V-slot of the input lacks a place node and it thus cannot be interpreted phonetically) because the clitic can be syllabified with the previous vocalic host, while in (65) epenthesizing the phonological features of the default vowel of Pallarese and linking them to the V-slot of the theme position solves the phonotactic problem (candidate a. thus wins over candidate b., which has an onset cluster *[lsp]). The reasons why this solution is optimal with respect to the epenthesis of a whole vowel (candidate c.) are the same as in the accusative plural, i.e, creating more structure (a new timing unit) is more costly than simply inserting phonological features and association lines into an already present structural position.

A different issue is at play regarding the deictic morph. As already mentioned, the motivations for the appearance of /i/ are still opaque and one can only conjecture about its distribution. In any case, I repeat once again the distribution of the deictic morph, emphasizing the compulsory appearance of /i/ in combinations, which will be the topic of discussion of sections 5.1.3.1 and 5.1.3.2:

- In the singular it always appears

- In the plural it does not appear in isolation

- *In combinations it always appears* (in both singular and plural)

### 5.1.3.1  The dative plural in combinations

Whoever tries to account for the phonological shape of the dative must deal with everything we have discussed thus far in this chapter. Phonology, morphology and syntax are surely intertwined in the final outcome of the clitic and no clear answer seems to be at hand. In fact, combinations with dative clitics, a good 'test' for the interfaces, are a puzzle in Pallarese Catalan. As we just saw, the deictic is not overt in isolation in the plural but it appears in combinations, which makes Pallarese different from other dialects like Standard Catalan or Central Catalan.

In Standard Catalan we find the dative forms *li-els*, where the deictic morph /i/ appears only in the singular, never in the plural:

(66)    a.    [**li**]        llegeixo un              conte
               DAT.3S read.1S a.MASC.S tale.MASC.S

'I read her/him a book'

b. [əlz]   llegeixo un      conte
  DAT.3PL read.1S a.MASC.S tale.MASC.S
  'I read them a book'

In combinations involving the plural the deictic never appears, yielding transparent outputs consisting of dative plus accusative forms (as opposed to Central Catalan or Pallarese, which generate opaque combinations with the deictic morph /i/). The table in (67) and the examples in (68) show that Standard Catalan always realizes the plural dative clitic as *els* when combined with accusative, genitive or neuter clitics:

(67)
| els | + | les | (acc.) | = | els les |
| els | + | els | (acc.) | = | els els |
| els | + | la | (acc.) | = | els la |
| els | + | el | (acc.) | = | els el |
| els | + | en | (gen.) | = | els en |
| els | + | ho | (neut.) | = | els ho |

(68)  a.  Les        joguines,  [əlz    ləs]          compro
       the.FEM.PL toy.FEM.PL DAT.3PL ACC.FEM.3PL buy.1S
       'I buy them the toys'

    b.  El         vestit,    [əlz    əl]          compro
       the.MASC.S dressMASC.S DAT.3PL ACC.MASC.3S buy.1S
       'I buy them the dress'

There is a total absence of the deictic in the plural in combinations (except when the dative plural clitic is specifically joined to the deictic-locative *hi*, as in *Els hi dono el llibre, a la platja* 'I give them the book there, at the beach', where *hi* refers to 'the beach') as well as in isolation, and, as just noted, this correlates with non-opaque clitic sequences.

It is clear from the examples and combinations above that there is an asymmetry between the singular and plural in Standard Catalan with regard to the appearance of /i/. Note that there is no such asymmetry *within* categories, i.e., within singular, where the deictic appears always, or within plural, where it is always absent.

In Central Catalan, by contrast, the deictic morph /i/ is always realized in the plural dative clitic, its form being *els hi* (plus the form *li* in the singular):

(69)  a.  [li]   llegeixo un      conte
       DAT.3S read.1S a.MASC.S tale.MASC.S

> *I read her/him a story*

    b.   **[əlzi]**    llegeixo   un          conte
            DAT.3PL  read.1S  a.MASC.S  tale.M.SG
            'I read them a story'

In combinations, Central Catalan generates opaque sequences because it is not possible to have a clear-cut segmentation of the clitic sequence. That is, while in (67) (Standard Catalan) the accusative and dative were distinguishable when put together, in (70) we see that this distinction becomes blurred in Central Catalan:

(70)

| els | + | les | (acc.) | = | els hi | [əlzi] |
|-----|---|-----|--------|---|--------|--------|
| els | + | els | (acc.) | = | els hi | [əlzi] |
| els | + | la | (acc.) | = | els hi | [əlzi] |
| els | + | el | (acc.) | = | els hi | [əlzi] |
| els | + | en | (gen.) | = | els en hi | [əlzəni] |
| els | + | ho | (neut.) | = | els hi | [əlzi] |

(71)    a.   Les          joguines,   **[əlz   i]**  compro
               the.FEM.PL  toy.FEM.PL  DAT.3PL DX  buy.1S
               'I buy them the toys'

        b.   El           vestit,     **[əlz   i]**  compro
               the.MASC.S  dressMASC.S  DAT.3PL DX  buy.1S
               'I buy them the dress'

In (71) the glosses connect the form *els* with the dative clitic and *hi* with deixis. However, these correspondences are not really so clear-cut, because in (71-a) the accusative antecedent is plural while in (71-b) it is singular. Is the accusative not spelled out in combinations? Does [z] realize the plurality of the dative or that of the accusative? What does /i/ stand for in combinations? Although the arguments of the verb are different in all cases, as depicted in (70), the final outcome is always the same, *els hi*, which in fact shares its form with the dative. These configurations are considered opaque because one cannot determine by their phonological shape a one-to-one morphological correspondence, whereas in Standard Catalan every argument of the verb is clearly transferred onto a specific clitic.

In any case, the fact is that /i/ is present in all instances of the dative in Central Catalan, be it singular or plural, in isolation or in combinations, thus posing no problems. There seems to be, thus, some kind of special property in the deictic marker in Central Catalan that makes it 'everlasting' in the expression of the dative (in all its dimensions).

In Pallarese, however, we find a different scenario. The situation differs from that of Central and Standard Catalan because there is an asymmetry *within* the plural. That is, in Standard Catalan, the morph /i/ appears only in the singular but never in the plural, being stable intracategorically, but in Pallarese there is no such stability in the plural. While in the singular its only shape is *li*, in the plural we find the alternation between *les* and *les hi*, with the deictic marker present only in combinations, as shown in (72) and (73):

(72)

| | | | | | | |
|---|---|---|---|---|---|---|
| les | + | les | (acc.) | = | les hi | [lezi] |
| les | + | la | (acc.) | = | les hi | [lezi] |
| les | + | lo | (acc.) | = | les hi | [lezi] |
| les | + | en | (gen.) | = | les hi'n | [lezin] |
| les | + | ho | (neut.) | = | les hi | [lezi] |

(73)  a.  Les         joguines,   **[lezi]**                    compro
          the.FEM.PL  toy.FEM.PL  DAT.3PL+ACC.FEM.PL  buy.1S
          *I buy them the toys*

  b.  El          vestit,   **[lezi]**                    compro
          the.MASC.S  dressMASC.S  DAT.3PL+ACC.MASC.S  buy.1S
          *I buy them the dress*

The puzzle of Pallarese dative plural clitics, then, lies in the realization of the deictic marker only in certain environments, namely in combinations. Why is Pallarese inconsistent in its use of the deictic? As noted above, I think the motivations for this behavior are both phonological and morphosyntactic.

As shown in (72), all combinations that involve the dative plural, no matter what the other argument of the verb is, end up being realized with the same form, [lezi] (except for the genitive + dative combination, which yields [lezin]).[17]  Although for some speakers these clitic clusters can be also expressed without the deictic, with the forms [ls]/[les] (e.g., *Lo lluç, **[les]** portes lo dilluns* 'S/he brings them the hake on Monday'), the tendency seems to point towards the adjunction of the deictic morph, as opposed to isolation contexts, where the tendency is reversed (some speakers use /i/ in isolated dative forms, but they are crucially the speakers who show most influence from Central Catalan). What makes combinations special with respect to the use of a deictic morph? While the reasons for this are obscure and

---

[17]These kind of clitic clusters are also phonologically conditioned, behaving as single plural datives or accusatives, e.g., *Diu **que'ls hi** [kelzi] duu lo dimarts, lo lluç* 'S/he says that s/he brings them the hake on Tuesday', where the clitic is attached to the previous vowel of the complementizer and thus does not need any vowel in the theme position for phonotactic reasons.

no general conclusions can be drawn from the data, it seems that using a specific
morph for combinations which are already opaque serves to signal its special status.
That is, as the arguments of the verb are not directly translated into a transparent
morphological template, with every case spelled out as a specific clitic, the use of
/i/ points out the 'abnormality' of the sequence.  In a way, it could be considered
a mark that helps the speaker/addressee identify the syntactic configuration of the
cluster.  For this reason, it could be said that, in combinations, the deictic serves a
morphosyntactic purpose, not a phonological one.

Therefore, it is understandable that the deictic appears again in combinations
with the dative plural. It is not clear, though, why these combinations are opaque.
Bonet (1991, 1995a) solves the problem of clitic clusters by introducing morpholog-
ical templates composed of different slots where the pronouns have to be inserted
(each pronoun consisting of a structured bundle of features).  In certain circum-
stances, the features of every different clitic will fight for the same slot, generating
conflicts that yield to the above mentioned opaque combinations.

With respect to dative + accusative combinations, the situation is the following.
As the 3rd person bears no [PERSON] feature, there is no slot for that feature in
the template of 3rd person clitic clusters.  Instead, there is one —only one— for
[ARGUMENT], which both the accusative and the dative share:

(74)     [ARGUMENT]   +   [GENITIVE]   +   [OBLIQUE]/[NEUTER]

Given the template in (74), just one of the the two features of either the acusative
or the dative can occupy the specific slot for [ARGUMENT], and therefore the
combination of dative and accusative clitics will generate the same output, whether
any of them be singular or plural.[18]  The fact that in dative singular + accusative
plural combinations we find the form [lezi] (or [əlzi] in Central Catalan), which
corresponds to a plural dative, not a singular one, clearly tells us that we are not
dealing with 'just' a dative clitic.  In that case, and assuming that the dative absorbs
the accusative, we would expect the singular dative clitic *li*. This is obviously not
the case. Note that in Bonet's analysis, the [OBLIQUE] feature, borne only by the
dative and locative, is spelled out as /i/, which appears last in the template. This
split between [ARGUMENT] and [OBLIQUE] in the structural configuration of the
dative also explains why /i/ appears after the plural morph (recall that [plural]
depends on the [ARGUMENT] node).

By contrast, Martín (2012) adopts a more syntactic approach to clitic clusters.

---

[18]There is no specific slot for [feminine] or [plural] in Bonet's analysis, as these do not consitute
primitive defining features of clitics.  As she clarifies, "these features, if present, are mapped
along with the node that dominates them." (Bonet, 1991, p. 102)

He argues that the fact that datives are already decomposable in subparts motivates their insertion when different clitics are put together. If accusatives are contained within datives and, furthermore, datives encode deixis, it is preferable to use the dative than the accusative clitic in combinations (which does not trigger dative interpretation). In fact, the accusative on its own can never replace a dative + accusative construction in Pallarese:

(75)  a.  Lo          regal,          als             nens,          **[lezi]**
          the.MASC.S present.MASC.S to+the.MASC.PL kid.MASC.PL DAT.3PL
          compro
          buy.1S
          'I buy them the present (for the kids)'

      b.  Lo          regal,          als             nens,          ***[lo]**
          the.MASC.S present.MASC.S to+the.MASC.PL kid.MASC.PL DAT.3PL
          compro
          buy.1S
          'I buy them the present (for the kids)'

In (75), the clitic that is chosen to replace the two verb arguments coincides with the dative plural form [lezi], although as already shown the deictic is usually absent in isolation. The fact that this form in (75-a) is preferred over the single accusative in (75-b) is due to the 'complexity' of the dative. Let me make this clear: while the accusative expresses only one argument of the verb and it is not possible to derive a dative interpretation from it, the structural composition of the dative works as a portmanteau for accusative and dative together. Martín's proposal for the syntactic structure of the dative relies on two DP layers, which explains, according to him, the choice of the dative clitic in third person combinations, and also why these structures can be opaque (though they need not be —in some dialects, like Standard Catalan and Valencian, they are transparent). If the accusative is generated in the lower DP and the dative interpretation is licensed only at $DP_2$, the dative clitic is naturally a better candidate to represent combinations of the two cases.

Although I agree with Martín in that the two DP layers proposed for the dative can express the complex nature of clusters, I believe that their sharing of the same form is a mere coincidence. That is, as pointed out by Bonet (1991), the use of the plural dative form in combinations where the accusative is plural and the dative is singular clearly demonstrates that we are not dealing with a 'true' dative. Thus, the form [lezi] (or [əlzi] in Central Catalan) can only be understood as a portmanteau, as in French preposition + article sequences.

(76)    à (preposition) + le (def. masc. sg. article) = au /o/

> *Je vais **au** parc* 'I'm going to the park'

It is clear that, in French, *au* encodes both the preposition and the definite article values, so neither of them are lost. We find the same phenomenon at work in Catalan clitic clusters. Departing from Bonet (1991), where the competition for the same slot motivated the association of only one of the two features in the morphological position, I assume that both features can be realized by the same morph. This happens, I argue, for both the plural and the definiteness morphs, /l/ and /z/, which always appear in dative + accusative combinations (plus the morph /i/, which corresponds to deixis):

(77)    l  (definiteness)  +  i  (deixis)  +  z  (plural)

These three different morphs appear in all combinations, no matter which element is plural in number, the accusative or the dative. If we assume that there is some kind of restriction on the co-occurrence of definiteness and plural morphemes (recall that this restriction is variable, as in other dialects two definiteness and two plural morphemes appear simultaneously), it seems clear why the final form is [lezi], a single clitic that contains all instances of the relevant features. I understand this restriction in terms of morphosyntactic dissimilation, as proposed in Nevins (2012). In phonology, OCP effects —i.e., the avoidance of consecutive identical features in underlying representations— are well established in the literature, and the idea is to treat the non-occurence of equal morphosyntactic features under the same view. Nevins (2012, p. 3) raises an interesting question regarding this morphosyntactic phenomenon:[19]

> The most interesting architectural question when facing haplology is, is the relevant OCP-violating sequence *generated but not realized* (e.g., due to deletion, coalescence or allomorphy), or are such sequences simply *never generated in the first place*?

Two issues are important here. First, are these particular illicit sequences ever generated in the syntax? And second, if they are generated at all, what is the repair strategy used to avoid them? In the case under investigation, it is clear that the syntactic structures are generated, because the clitics are coindexed with otherwise fully realized arguments of the verb. That is, in non-clitic environments, accusatives and datives co-occur in the same sentence. With regard to the second question, things get more complicated. As we saw recently, Bonet (1991) argues that only

---

[19]Haplology or OCP can be used interchangeably to refer to this particular idea.

one of the [ARGUMENT] features of either the dative or the accusative clitic can occupy the slot assigned for it, which implies that deletion, one of the three options provided by Nevins to solve morphosyntactic OCP, is the preferred repair for Bonet. I argue, instead, that coalescence better captures the nature of clitic combinations.

It is not the purpose of this thesis to provide a full analysis of clitic combinations in Catalan and other Romance languages, which still today remain a mystery in many respects. However, it is clear from the outputs shown in (72) that something must be going on with combinations of 3rd person clitics. Departing from the idea of morphological dissimilation, the opaque outputs can be understood by the morphological constraint in (78):[20]

(78)  *[+def][+def]

Both dative and accusative *l*-clitics express definiteness, and (78) goes against the multiple expression of this feature (in clitic combinations). Deleting one of their instances is a possible solution, but not the only one, since morphological coalescence solves the problem as well. Coalescence strikes me as the most plausible solution here, as it has been reported to be by Heath (1998) for similar cases in Native American languages, where the portmanteau is used for a sequence of 1st and 2nd pronouns instead of concatenating their forms in isolation.

In Catalan, however, the exponence of the portmanteau resorts to the morphs that conform the isolated clitics without the need to completely change their phonological shape (as in French preposition + article sequences). The fusion of the two [+def] features in one morpheme avoids the constraint in (78) but creates, at the same time, an opaque form, as in (80):

(79)   *Independent realization*          (80)   *Coalescence (opaque form)*
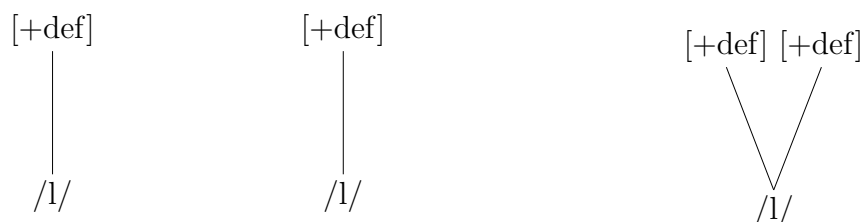


The OCP morphological restriction of (78) is not respected in (79), and thus the coalescence solution is preferred.[21] The same happens with the plural morpheme: if the two arguments of the verb are plural, only one morph realizes both of them:

---

[20]Constraint must be understood here as 'filter'. I am not arguing that this is an OT constraint.

[21]The situation depicted in (79) is what we find in dialects that tolerate sequences of dative + accusative clitics, like Standard Catalan (e.g., *els el*, dat. pl. + acc. sg.).

(81)
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| les | (dat) | + | la | (acc.) | = | les hi | [lezi] |
| les | (dat) | + | lo | (acc.) | = | les hi | [lezi] |
| **les** | **(dat.)** | **+** | **les** | **(acc.)** | **=** | **les hi** | **[lezi]** |

In (81), the effects of this fusion process are made clear. The plural dative and plural accusative, both realized as *les*, end up taking the form *les hi*, with only one instance of [+def] and one instance of [+pl]. In the latter case, though, the problem could be morphophonological, not morphosyntactic, as the contact between two plural morphs would generate an illicit phonological string */zz/.

It is also worth mentioning that the Impoverishment rule that deleted gender features in the presence of [+plural] and [+deixis], repeated in (82), guarantees that gender is never realized even when one of the accusative arguments is singular:

$$(82) \qquad [\pm\text{fem}] \longrightarrow \emptyset / [+\text{def}] \underline{\quad} \left\{ \begin{array}{c} [+\text{pl}] \\ [+\mathbf{dx}] \end{array} \right\}$$

That is, the rule does not apply when it faces a singular accusative clitic because no plural or deictic features license it. When clitics are combined, though, the target features of the Impoverishment rule are met, as there is a feature bundle consisting of dative plus accusative feature subsets. Therefore, (82) correctly predicts that, in combinations, gender will never be realized, which I think is a neat result of this proposal. Be that as it may, one cannot know by empirical means which of the two options proposed in the literature, deletion or coalescence of identical features, is the most adequate for explaining the behavior of Catalan clitics. Intuitively, though, it seems more plausible to me that both accusative and dative are expressed within the same clitic output.

There are two remaining combinations that we have thus far not analyzed: dative plural + genitive and dative plural + neuter. The prediction made by Bonet (1991)'s template for Barceloní, shown in (74), is that the genitive case will appear in between the argument clitic, which we can identify with the morph /l/, and the oblique feature, which corresponds to /i/ in Bonet's proposal. This is actually what we find in Barceloní in forms like [əlzəni], where the partitive clitic *en* is intercalated between [ARGUMENT] and [OBLIQUE]. However, the analogous form in Pallarese is [lezin], going against the Barceloní template. If, as mentioned above, the marker /i/ appears in dative + accusative combinations to signal the opacity of the form, we should not expect it in dative plural + genitive combinations because the output is transparent. Therefore, the form should not contain /i/, which was found to be true for one speaker:

(83)  As        teues      germanes,  dóna'  **[lzne]**      tres, de
      to+theFEM.PL your.FEM.PL sister.FEM.PL give.IMP DAT.PL+GEN three of
      roses!
      rose.FEM.PL
      *Give your sisters three roses*

In (83), the form is [lzne] (which is phonologically conditioned and has no theme vowel -*e*), with no trace of the deictic. The dative clitic appears, thus, in its usual form in isolation followed by the clitic *ne* (we will see in chapter 6 that the vowel -*e* of the clitic *ne* is actually thematic as well). For this set of speakers, then, there is no problem regarding this combination, which consists in the juxtaposition of two independent clitic forms. But why is the form [lezin] the most common? We might relate the use of the morph /i/ in these speakers with the idea of Paradigm Uniformity or Uniform Exponence (Kenstowicz, 1996), in that using the form [lezi] plus the adjunction of the genitive [n] generates a regular paradigm for the whole plural dative in combinations (with either accusative or genitive). The same idea can be argued to rule out a form *[lez.ni], with the deictic after the genitive clitic. It is important to point out that this is the case for the previously mentioned Central Catalan form [əlzəni], which, according to Martín (2012), respects the case sequence established in Caha (2009, p. 10), shown in:

(84)    Nom  >  Acc  >  Gen  >  Dat  >  Instr  >  Com(itative)

The form [əlzəni] reflects the hierarchy established above, as the genitive clitic *en* is found between the accusative and the deictic part of the dative clitic. By contrast, Pallarese does not respect the case sequence, probably because of the paradigmatic effects just mentioned. In any case, what is clear is that we could never find a form *[lezn], with only one phonologically conditioned theme vowel in the accusative, because it would be banned by the phonotactics of Catalan. For this reason, only a form [lezne] (or [lzne] when there is a previous vocalic host) is allowed, with two phonologically conditioned theme vowels, in both the accusative and genitive clitics. The form [lezin] also allows the sillabyfication of the genitive due to its attachment to the deictic /i/, and thus the non-vocalic form [n] surfaces, so it is somehow phonologically preferable in this respect (no need to insert a new vowel), besides the fact that it maintains uniformity within the paradigm.

Regarding the combination with the neuter, the solution proposed in Bonet (1991) involves delinking certain features from the morphosyntactic structure of clitics and thereby yielding morphological outputs with the same featural composition of the dative, which causes the insertion of its corresponding phonological expo-

nents. In plural dative + neuter combinations, then, a morphological rule delinks the node [NEUTER], connected to [ARGUMENT], and thus the final output consists of [ARGUMENT] and [OBLIQUE] features, precisely those that conform the dative.

Martín (2012), on the other hand, argues that there is clitic dropping of the neuter *ho* because the dative does the work of both clitics, precisely because it is complex in nature (i.e., it already contains an accusative). However, if the coalescence hypothesis defended above for the accusative + dative combinations is correct, whereby two [+def] features could not be morphologically consecutive, the same can be expected in neuter + dative. The neuter takes a different shape than *l*-accusative clitics, but [+def] is also present in its featural composition (the presence of [+def] is confirmed by Spanish, a language that does not have a specific clitic for the neuter argument and uses the form *lo*, shared with the masculine accusative clitic).[22] The same solution as in (80) will be favored by the grammar, choosing an /l/ exponent that almost always characterizes the expression of definiteness.

### 5.1.3.2   The dative singular in combinations

The situation in combinations with the dative singular does not differ from what was shown above for the dative plural. The outcomes are, again, opaque clusters (except for the genitive), as shown in (85):

(85)

| | | | | | | |
|---|---|---|---|---|---|---|
| li | + | les | (acc.) | = | les hi | [lezi] |
| li | + | la | (acc.) | = | li | [li] |
| li | + | lo | (acc.) | = | li | [li] |
| li | + | en | (gen.) | = | li'n | [lin] |
| li | + | ho | (neut.) | = | li | [li] |

Before, it was noted that the fusion of the dative singular with the accusative plural yields [lezi], a form that coincides with the dative plural. As the dative is singular in number, it is impossible to treat [lezi] as the sole exponent of the dative. Again, the concept of portmanteau is what best defines the forms in (85).

A welcome result of the proposal has to do with the non-realization of gender in these combinations, as already noted in the previous section. In the plural, both for datives and accusatives, Pallarese never shows gender differences, but in the singular there is a contrast between masculine and feminine (in the accusative clitics

---

[22]The neuter clitic refers to arguments of the verb as well, although only to objects, not persons like *l*-accusative clitics. Gender is thus never expressed in it.

only). I proposed, however, an Impoverishment rule that deleted gender features in the presence not only of [+pl] but also of [+dx]. In combinations with the dative singular, the features of both the accusative and the dative are put together, which implies, once again, that the gender features which are the target of the rule in (82) are present in the accusative singular as well. This rule plus the feature co-occurrence restriction that disallows the presence of two consecutive [+def] features leaves no other possibility than realizing the combinations of dative singular + accusative singular as [li].

The combination *li + ho* does not pose any problem because it patterns like its plural counterpart, discussed above (the only difference involves number, but the morphological OCP constraint that bans consecutive [+def] features also operates in the singular).[23] As for *li + en*, there is a crucial difference with respect to the Barceloní dialect studied in Bonet (1991). The combination in Barceloní is not transparent, because the fusion of dative + genitive yields *ni*. The *l*-morph thus disappears, again due to delinking of the node [ARGUMENT] of the 3rd person dative clitic, which leaves it only with the oblique feature corresponding to /i/ (and thus the output *ni* when it is joined to the genitive). The situation in Pallarese requires no special comment because it behaves exactly as expected. Moreover, it creates an optimal paradigm, as in both singular and plural all the combinations become [li] and [lezi], except for the genitives [lin] and [lezin], which still follow the same pattern.

All in all, the data analyzed in this section regarding dative clitics and their relation to the accusative show that:

- The deictic is part of the dative but an independent unit as well, not used equally in all contexts.

- Morphological repairs are used in inflectional positions in order to avoid morphophonological mismatches. If an epenthetic site is not morphologically salient (i.e., not inflectional), a purely phonological repair is chosen. If, by contrast, it is salient, a morphological solution is preferred, as in the use of the theme position of the dative clitic that licenses the insertion of the features of the default vowel of Pallarese, *-e*.

---

[23]For some speakers this combination can be spelled out as [laj], as already mentioned in chapter 2. This output is truly opaque and I could find no solution for it from either a morphosyntactic or a phonological point of view.

## 5.1.4 Non-phonologically conditioned variation in the plural dative/accusative

An issue which has thus far received no attention is the variation found in the shape of plural dative and accusative clitics as described in 2.2. In fact, four different exponents can be found for the same clitic, be it accusative or dative (or definite article, as well):

(86)    *les    es*

       *ls    s*

We have seen throughout this chapter that the forms *ls/les* are phonologically conditioned, *les* appearing in absolute phrase-initial position or between consonants and *ls* after a vowel that allows its syllabification as a coda cluster. However, as shown in (87), (88) and (89) (for clarification purposes, some of the examples given in 2.2 are repeated here), the forms *es* and *s*, which have lost in both cases the definiteness morph *-l* —plus the theme vowel in the *s* variant— can appear in different contexts:

(87)    *Dative*

    a.  [**les**]    toques    la    mà
        DAT.3PL touch.2S the.FEM.S hand.FEM.S
        'You touch their hand'

    b.  Cosís-    [**ez**]    lo    botó!
        sew.IMPR DAT.3PL the.MASC.S button.MASC.S
        ´Sew the button for them!'

(88)    *Accusative*

    a.  [**es**]    tomates    [**es**]/[**les**]    fiques a l'
        the.FEM.PL tomato.FEM.PL ACC.3PL.FEM put.2S to the.MASC.S
        amanit
        salad
        'You put the tomatoes in the salad'

    b.  [**es**]    tomates    no' [**s**]    comprem ací
        the.FEM.PL tomato.FEM.PL NEG ACC.3PL.FEM buy.1PL here
        'We don't buy the tomatoes here'

(89)    *Definite article*

    a.  [**lez**]    més petits    van    a Sort i    [**ez**]
        the.MASC.PL more little.MASC.PL go.3PL to Sort and the..PL
        altres    a Alins
        other.MASC.PL to Alins
        'The younger ones go to Sort and the others to Alins'

    b.    **[es]**      cases      **[pes]**      que  hi  visquívom
          the.FEM.PL house.FEM.PL for+MASC.PL that LOC live.IMPF.1PL
          'The houses for those of us who lived there'

Given the data above, we may postulate the following:

1. *es* and *s* tend to appear, respectively, in the same phonological contexts as *les* and *ls*.

2. The variants without the *l*-morph occur freely with respect to their *l*-counterparts, without any apparent motivation for the choice of one over the other (speakers can use different forms indiscriminately in the same sentence, as in (88-a)).

It is logical to assume, then, that these alternations constitute a case of free variation. Bonet & Harbour (2012) argue that free variation can be regarded as a form of allomorphy in which no contextual information is sufficient to determine which variant should be used. The examples provided there are the Spanish Imperfect Subjunctive (*cantara/cantase* 'that he sang') and the Catalan word for 'nothing', *res* (which can be phonetically [rɛ́] or [rɛ́s]), and in both cases there is no specific reason that favors one allomorph over the other (assuming that two forms are listed in each lexical entry). While I agree with Bonet & Harbour (2012)'s idea, it cannot be extended to Pallarese because no allomorphy is assumed for 3rd person plural clitics (and definite articles). There is one single underlying form /l–V+z/ , where the theme position is filled with an underspecified vowel that gets its place features epenthesized under certain conditions. The fact that this vowel appears in the variant *es* even when it seems that it is not needed for phonotactic purposes reinforces the proposal advanced in this thesis about the theme position and goes against a purely epenthetic analysis.

Recall that we are assuming a structure for the plural clitic as follows:

(90)
```
              #
            /  \
          D      #
        /   \     |
       D    Th    z
       |    |
       l    e
```
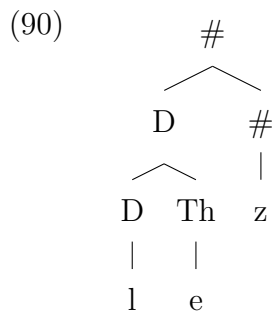
In (90) we see the morphological structure for the form /lez/, where the features of the default vowel of Pallarese, *-e*, have been epenthesized. If, for reasons that are not clear, the speaker does not realize the definiteness morph *-l*, the theme position

with its associated vowel still remains in the structural configuration of the clitic.[24]
Thus, the fact that *-e* occupies a morphological position makes it more 'resistant',
which allows its appearance without the defining exponent of the clitic. Without
this inflectional position, a purely epenthetic solution (i.e., regarding the vowel of
the clitic *es* as non-morphological), which is more dependent on the phonetics component, seems less motivated.

If the underlying form is /l–V+z/, the derivation of all the Pallarese variants
could be as shown in (91):

(91)    /l–V+z/  >  les  >  es  >  s
        /l–V+z/  >  ls   >  s

We see in (91) that *es* and *s* could be generated in different ways. While it is clear
that *es* is derived directly from the form *les*, the generation of the *s* variant could
come from either *es* or *ls*. There is a tendency to delete *-l* before *-s*, made clear
in the portmanteau *pes* (preposition *per* + definite article; sg. *pel*) in (89-b), so
it seems plausible to derive it from *ls*. We should not discard the other option,
however, whereby it could be derived from *es* by the deletion of the theme vowel.
Note, however, that it is not possible to find the *s* clitic in absolute phrase-initial
position, even if it would be phonotactically acceptable:

(92)   *[**z**]         arbres        són      grans
       the.DEF.PL tree.MASC.PL be.3PL big.MASC.PL
       'The trees are big'

The data in (92) seem to indicate, then, that the direction of the derivation for the
*s* variant is /lz/ > s, because only *les* and *es* can appear in this position, never [lz]
—due to phonotactic reasons. If *es* is found phrase-initially and *s* were derived from
*es*, we should expect sentences like those in (92). As this is not the case in Pallarese,
we are forced to conclude that only the derivation from *ls* is plausible.

Nevertheless, the behavior of these plural clitics is sometimes unexpected, as
in (89-a). After the conjunction *i* it would seem to make more sense to use *s*, as
in (88-b), because this would prevent the two vowels *-i* and *-e* from coming into
contact, but *es* is preferred instead, even when the tendency of the speakers who
were interviewed is to use *s* in these contexts —and this seems reasonable, given

---

[24]One could hypothesize that a constraint like REALIZE-MORPHEME (van Oostendorp, 2005)
–"for every morpheme in the input, some phonological element should be present in the output"–
plays a role in the presence or absence of the definiteness morph, but the problem lies in the
variability of the realization process. Perhaps Stochastic OT (Boersma, 1997) could model the
variation found in Pallarese, but any attempt to formally implement this idea would fall well
outside the scope of this dissertation.

that they systematically use the non-vocalic form [ls]/[lz] when there is a preceding vowel.

It is also puzzling that the whole clitic structure is realized by the plural morph /z/. That is, either the plural accusative, dative or definite article are expressed by one single morph which encodes plurality. As already noted in previous sections, there is something about number that makes it special with respect to other categories. First, gender was neutralized in the presence of [+plural], and now even the definiteness morph, the 'central' exponent of the clitic, is not realized, something that never occurs in the singular. Again, why number should trigger all these changes is a difficult question to answer but nonetheless a highly intriguing one.

All in all, we can sum up the conclusions drawn from the data as follows:

1. The forms *les* and *ls* of the plural clitic appear in complementary distribution, as determined by the phonological context imposed by their host.

2. The forms *es* and *s*, which can replace *les* and *ls*, respectively, are in free variation, with speakers indiscriminately using the *l*-clitics or their non-*l* counterparts.

# SIX

# OTHER PRONOMINAL CLITICS

As we saw in section 5.1, Pallarese Catalan 3rd person clitics differ from 1st and 2nd person clitics in their morphosyntactic composition, especially with regard to gender marking. This chapter is devoted to the analysis of 1st and 2nd person clitics, as well as the rest of the clitics –i.e., reflexive, locative, neuter and partitive–, focussing on their morphological structure and the effect it has on their final phonological shape.

While 3rd person can equate to 'no person', 1st and 2nd persons are the active participants of the discourse. Speaker and addressee must be encoded in 1st and 2nd person pronouns, respectively, and, therefore, reflected in their featural composition. In this chapter I will not get into the details of the featural specifications of clitics, as whether we assume the hierarchy for person features found in Bonet (1991) or Heap (2005), for example, is not relevant for our present purposes. Instead, I will discuss here the role of the morphological theme position assumed in this thesis and the implications it has for the phonological component.

In 6.1 I analyze 1st and 2nd person singular clitics, in addition to the partitive and the reflexive, and in 6.2 I analyze their plural counterparts. Finally, the neuter and the locative are presented in 6.3.

## 6.1   1st and 2nd person singular, partitive and reflexive

1st and 2nd person singular clitics have a similar shape, differing only in the morph that expresses 1st person, /m/, and 2nd person, /t/. Apart from this, they behave identically at the prosodic level and they follow the same morphological ordering stated in (25) of chapter 2, i.e., after the reflexive but before 3rd person accusative

and the partitive/locative.[1] In turn, the partitive (*ne*) and reflexive (*se*) clitics are identical to 1st and 2nd person clitics in terms of their phonological and morphological structure. The only difference is, obviously, their featural composition, which consists of [+genitive] for *ne* and [+reflexive] for *se*.[2]

Table 6.1 shows the phonetic outputs of 1st and 2nd person clitics, partitives and reflexives, where we see that the only vowel that can appear at their right edge is -*e*:

| | | |
|---|---|---|
| **1st person** | [m] | [me] |
| **2nd person** | [t] | [te] |
| **partitive** | [n] | [ne] |
| **reflexive** | [s] | [se] |

Table 6.1: *Clitics' phonetic forms*

The forms shown above are phonologically conditioned, but their underlying representation has just one morph in all cases, /m/, /t/, /n/ or /s/. When these morphs can be attached to a vowel, as in (1), they appear in their non-vocalic form; when they cannot, they surface with the vowel -*e*, as in (2):

(1)  a.  Porta'      [**m**]    lo          sopar
         bring.IMP2S.  DAT.1S  DEF.MASC.S  dinner.MASC.S
         'Bring me the dinner!'

     b.  Ja'    [**t**]    poden  dir      açò
         already  DAT.2S  can.1PL  say.INF  this.NEUT
         'They can tell you this'

     c.  [**n**]'  agafa    quatre
         PART  take.3S  four
         'S/he takes four of those'

     d.  No  [**s**]'  ha            trencat
         NEG  REFL   have.AUX.3S  broken.PART
         '(This) didn't break'

---

(2)   a.   [**te**]   vull   dir   açò
           DAT.2S want.1S say.INF this.NEUT
           'I want to tell you this'

      b.   [**me**]   vai   prejubilar
           REFL.1S go.1S pre-retire.INF
           'I got pre-retired'

      c.   D' aguest   [**ne**]   vull   tres
           of this.M.S PART want.1S three
           'I want three of this'

      d.   [**se**]   fa   aquí
           REFL do.3S here
           '(This) is done here'

The distribution of these phonologically conditioned forms is shown below:

| _C | C_C | _V | V_V | C_V | V_C |
|------|------|------|------|------|------|
| [me] | [me] | [m] | [m] | [m] | [m] |
| [te] | [te] | [t] | [t] | [t] | [t] |
| [se] | [se] | [s] | [s] | [s] | [s] |
| [ne] | [ne] | [n] | [n] | [n] | [n] |

With regard to their morphological structure, Pallarese 1st and 2nd person pronominal, genitive and reflexive clitics parallel the 3rd person clitic structure that we saw in example (1) of 5.1 –which includes a theme position, like nouns and adjectives–, repeated below for clarity:

(3)   *3rd person pronominal accusative clitic*

```
            D
          ⌒
       D     Th
       |      |
       l    Ø/o/a
```
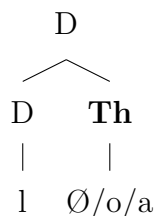
In (3), it is shown that the D head of the clitic is the definiteness morph /l/, which has /o/ and /a/ as possible theme vowels, depending on whether they are the exponents for masculine or feminine, respectively. We have seen in chapter 5 how the features associated with gender are only linked to the theme position under certain conditions, and this is why we can also find the 3rd person clitic with no features associated with the theme position, thus being phonetically empty, such as [l]. By contrast, genitive, reflexive, 1st and 2nd person pronominal clitics lack any gender

whatsoever and, as they cannot get the features associated with /a/ (feminine) and /o/ (masculine) in any case, we only find the alternation *-Ø/-e* (hence the difference in the phonologicall spell-out of the theme positions in (3) and (4)):

(4)     *Genitive, reflexive, 1st and 2nd pronominal clitics (accusative and dative)*

<center>

D

D    **Th**

|      |

m/t/n/s  Ø/e

</center>
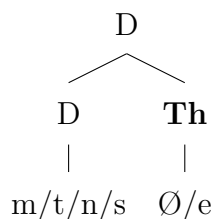
The D-head is occupied by /m/,/t/, /n/ or /s/ in (4), and the theme position is filled by either *-e* or a phonetic *-Ø*. Even if gender is not part of the featural representation of these clitics, they maintain the morphologically salient theme position, as already happened with 3rd person plural clitics that only show the alternation *ls/les*. In fact, the aim of this section is to show that, as we saw for 3rd person plural clitics, which lacked gender altogether due to Impoverishment, the vowel *-e* found in genitive, reflexive, 1st and 2nd person clitics is not entirely epenthetic. It is actually a theme vowel even though its features are epenthesized, something that I will argue for below as the only option available to repair syllable structure is the insertion of the features associated with *-e*, the unmarked vowel of Pallarese, as the features associated with gender are not available due to its absence in the featural specification of these clitics.

Traditionally, the alternation between vocalic (those that surface with *-e*) and non-vocalic clitics in Northwestern Catalan has been interpreted as a product of epenthesis (e.g., Massanell 2011, Clua et al. 2013). That is, when there is a syllabification problem because the clitic cannot be attached to a vocalic host, as in the examples of (2), an epenthetic vowel is inserted. This could be an option but, as shown in previous chapters, epenthesis is avoided at the right edge of the word due to the morphological salience of the theme position.

I said before that the underlying representation of 1st, 2nd, partitive and reflexive clitics is /m/, /t/, /n/ and /s/, but, to be more accurate, we have to assume that it is in fact /m-V/, /t-V/, /n-V/ and /s-V/. The same process of phonological spell-out of 3rd person clitics holds for these other clitics, with the theme postion being filled by an underspecified vowel without place features, as we saw previously in other chapters, repeated in (5):

(5)     Underspecified vowel in the theme position:
        Th ↔ V (*à la* Roca 2010)

This is the regular process of theme vowel assignment, which guarantees the presence of a vowel in the theme position but does not specifiy its place features. Any element of the nominal system will generate a theme position, as will these clitics. What differs from 3rd person clitics is the second step in their final phonological spell-out, i.e., the insertion, in this case, of the place features associated with gender, shown again in (6):

(6)     Association of gender values with phonological features:

|  |  |  |  |
|---|---|---|---|
| [+fem] | ↔ | [+low] | (= /a/) |
| [–fem] | ↔ | [+labial, –high] | (= /o/) |

This spell-out step cannot be implemented because reflexive, partitive, 1st and 2nd person clitics lack gender features, as opposed to 3rd person clitics. If there are no gender values, then, it is impossible to insert the phonological features associated with them. This is exactly what happened in the case of the 3rd person plural clitic, *ls/les*. As we saw in 5.1.2, there is a process of Impoverishment in the context of plural + definiteness, and the gender features that trigger the regular feminine or masculine spell-out are deleted. Here, instead, gender does not need to be deleted because it was never part of the featural composition of these clitics, but the parallelism with the 3rd person plural clitic is clear, as in both cases the chosen vowel is *-e* in the same phonological environments, i.e., when the clitic cannot be attached to a vocalic host. The arguments that were advocated in the case of the 3rd person plural clitic, regarding vowel quality, hold for these clitics as well, and so does the constraint ranking. Thus, we can see in tableaux (8) and (10), exemplified with the partitive clitic /n/, but applicable to all the clitics mentioned above, the generation of both output forms as determined by the ordering of the constraints:

(7)     [**ne**]   porto
        PART   bring.1S
        *I bring some of these*

(8)     *Consonantal host (phonotactic problem)*

| | /n  **V**  #  p  ɔ  r  t  o/ | DEP-V | σSTR | DEP[+HI] | DEP[-BK] | DEP-LK |
|---|---|---|---|---|---|---|
| a. ☞ | [-back]<br><br>n  e      p  ɔ́  r  t  o | | | | 1 | 1 |
| b. | [+high]<br><br>n  i      p  ɔ́  r  t  o | | | 1W | L | 1 |
| c. | n      p  ɔ́  r  t  o | | 1W | | L | L |
| d. | [-back]<br><br>e  n      p  ɔ́  r  t  o | 1W | | | 1 | 1 |

In (8) it is shown that the pattern of the 3rd person plural clitic is repeated. There are no available floating features in the input as there is no gender but, nevertheless, something needs to be inserted in the output in order to correctly sillabify the clitic, as the faithfulmost candidate, candidate c., is ruled out by σSTR (falling sonority onset clusters are not allowed in Catalan). Candidate d. is dispreferred because, as in the rest of the clitics analyzed in previous chapters, epenthesizing place features, an X-slot plus association lines is more costly than just inserting place features, the option taken by candidates a. and b. Between these two, a. is the winner due to the greater adequacy of the place features of *-e* over those of *-i*, as was already argued in chapter 5 (in this context, *-e* is preferred over *-i* as an epenthetic vowel due to perceptual salience and positional markedness).

The situation changes when there is a vocalic host for the clitic, as in the 3rd person plural clitic:

(9)  no  [**n**]  porto
     NEG PART bring.1S
     'I don't bring some of these'

(10)  *Vocalic host*

| /n o # n **V** # p ɔ r t o/ | DEP-V | σSTR | DEP[+HI] | DEP[-BK] | DEP-LK |
|---|---|---|---|---|---|
| a. ☞  n  o    n    p  ɔ́  r  t  o | | | | | |
| [+high] ‖ b.  n  o    n  i    p  ɔ́  r  t  o | | | 1W | | 1 |
| [-back] ‖ c.  n  o    n  e    p  ɔ́  r  t  o | | | | 1W | 1 |
| [-back] ‖ d.  n  o    e  n    p  ɔ́  r  t  o | 1W | | | 1 | 1 |

In (10), any candidate that epenthesizes association lines or features is harmonically bounded by candidate a., which is simply the most faithful candidate and does not violate any of the constraints at play that regulate phonotactics or vowel insertion. The representation of candidate a. can be misleading because it creates the impression that the V-slot (i.e., a [–cons] feature and a skeletal position) has been deleted. This is what we can infer from (10) but, in fact, the underspecified vowel appears in the output, although it will not be interpreted phonetically due to Stray Erasure (it does not appear in the tableau in order to exemplify a clearer opposition between candidates with a full vowel and candidates without it).

The novel treatment of these alternating vocalic and non-vocalic clitics relies on their morphological affiliation. Previously, these vowels have been considered to be purely epenthetic, but this thesis shows that their location favors a thematic interpretation. The features of the vowel are epenthesized but, crucially, the V-slot of the functional head's theme position attracts those features and the whole vowel thus becomes thematic. An argument in favor of this idea of a morphologically salient position comes from examples where -*e* is maintained in the clitic even when it is not needed for syllabification, as shown in (11):

184

(11)    No  [**me**]   compra açò
        NEG DAT.1S buy.3S  this.NEUT
        *S/he doesn't buy this to me*

It was previously argued in this section that a form like (11) should not exist because, in V_C contexts, we find the non-vocalic form. While this is true in the majority of cases, where [m] (or [t], [s] or [n]) can be attached to a preceding vowel, sometimes it is also possible to find the full form even when it is not needed for syllabification purposes. This implies that these forms must be stored somehow, and the salience of the morphological position seems to provide a better explanation for it. That is, pure epenthesis is a phonetics-phonology interface phenomenon which is more dependent on articulatory or perceptual factors, whereas a theme position is linked to the morphological structure of the nominal system as a whole. Thus, it is more stable and, therefore, more easily stored with a lexical entry.

## 6.1.1    Combinations

The combination of *me*, *te*, *se* and *ne* with other clitics follows the general morphological restrictions imposed on Catalan clitics. They can be grouped together according to these schemata:[3]

1.  *se + me, te, ne, mos, vos*

2.  *me, te, se + ne, ho*

3.  *me, te, se, ne + hi*

Other combinations and orders of these clitics are not possible.[4] This is what happens, for example, in impersonal + reflexive sequences. According to Bonet (1991), sentences that contain these two consecutive clitics are not possible due to the limitations of the morphological template that the cluster has to adapt to:[5]

(12)    *Es es    renta
        IMP REFL wash.3S
        'One washes oneself'

---

[3]The combination of some of these clitics with 3rd person pronouns has been already analyzed in section 5.1.3 of chapter 5.

[4]Some other combinations that involve the neuter clitic *ho* –which are not mentioned here– are analyzed in section 6.3. Clusters involving the 1st and 2nd person plural clitics are discussed in 6.2.1.

[5]Pescarini (2010) treats, instead, the impossibility of (12) as an instance of an OCP-like morphological constraint.

As, in her analysis, there is only one slot for the [PERSON] clitic, one of the two, the reflexive, "has to be spelled out in the foot of the chain it belongs to" (Bonet, 1991, p. 106):

(13)  Es  renta  **a un mateix**
      IMP wash.3S to one self.MASC.S
      'One washes oneself'

The morphosyntactic conditionings on clitic clusters are (mostly) not cosnidered in this dissertation. Instead, I will focus here on the phonological spell-out of the combinations that are in fact possible in Pallarese.

Within the first group, we mostly find combinations in which the second clitic appears as a non-vocalic form, as in (14):[6]

(14)  [**se  n**]  surt
      REFL GEN leave.3S.FUT
      *S/he manages well*

The result is the same with 1st and 2nd person singular pronouns, i.e., [sem] and [set]. If we assume, as I am doing, that the clitics in (14) are underlyingly /s/ and /n/, there is a vowel in between whose appearance needs to be accounted for. The fact that every clitic projects a theme position leads us to determine that this vowel is in fact morphological, as in the isolated clitics. In principle, if one does not posit a theme position, the vowel, which must be epenthetic, could be linked to either /s/ or /n/. If, by contrast, we assume that every clitic projects a theme position, the vowel must be linked to the first clitic, *se* in this case. The tableau in (15) shows this:[7]

---

[6]I already mentioned that speakers sometimes use the full form of the clitics in isolation even when the phonological environment provides the context for their realization as non-vocalic, as in *que [se] parla* 'that it is said' (we would expect *que* [s] *parla* given that the clitic could be attached to the previous vowel). The same happens in clitic clusters, where we find forms such as *[se te] moriva el videll* 'that your calf died', although they are clearly much less frequent.

[7]There is coda cluster simplification in Catalan and this is the reason why /surt/ surfaces as [súr] in (15).

(15)   [**sen**] surt

| /s **V** # n **V** # s u r t/ | Dep-V | σStr | Dep[+hi] | Dep[-bk] | Dep-Lk |
|---|---|---|---|---|---|
| a. ☞  [-back] over **s e** n s ú r | | | | 1 | 1 |
| b.  s  n  s ú r | | 1W | | L | L |
| c.  [-back] [-back] over **s e** n **e** s ú r | | | | 2W | 2 |
| d.  [-back] over s **e** n s ú r | 1W | | | 1 | 1 |
| e.  [+high] over **s i** n s ú r | | | 1W | L | 1 |
| f.  [-back] over s n **e** s ú r | | 1W | | 1 | 1 |

In the input, the theme position is filled with an underspecified vowel, a general process in the whole nominal system. This vowel can only be interpreted phonetically if it has place features, which is precisely what is absent from the input. Candidate b. is the most faithful candidate because it lacks those features in the output as well, but it is ruled out due to the violation of well-formedness syllable conditions. A vowel is needed in order to syllabify the cluster, and the candidate that needs to insert less material is the winning candidate, i.e., candidate a. Whereas candidate d. needs to insert a whole peripheral vowel that belongs to the genitive /n/, which has no connections with the morphological structure of the clitic, candidate a. only needs to epenthesize the features corresponding to the vowel *-e* and the association lines that link them to the skeleton (the V-node created by the theme position). Therefore, it is preferred over candidate d. and also over candidate c., because c. epenthesizes features and association lines in the theme positions of both clitics, something which is, in principle, not needed for syllabification. Candidate f. shows,

moreover, that if there is feature epenthesis it must be in the theme position of the first clitic, as doing so in the second one does not solve the syllabification problem. Finally, candidate e. is ruled out due to its vowel quality, as we already established that *-e* is the unmarked epenthetic vowel for Pallarese Catalan.

According to the morphological structure argued for in this thesis, candidate a. is the only possible outcome when these two clitics appear in combination. That is, the nature of the theme position, which is morphologically salient, bans the use of a purely epenthetic vowel at the right edge of the word (i.e., the location for inflection) to avoid a mismatch between phonological and morphological structure. Instead, the proposed autosegmental lexical insertion process predicts that, in case of phonological conflict, the V-slot of the theme position will favor the insertion of the default features of the unmarked vowel of Pallarese; full epenthesis (i.e., the creation of a skeletal position plus the insertion of features) is a last resort strategy. This is, in fact, a felicitous prediction of the proposal, as, contrary to purely epenthetic analyses, it can determine the exact affiliation of the vowel *-e* in clitic combinations (if a morphological position of some kind is not posited, it could both belong to the reflexive or the genitive clitics, i.e., it could be a left or right peripheral vowel).

The rest of the combinations follow the same exact pattern exemplified in (15), with the first clitic of the cluster filling up the theme position. There are only two comments that need to be made with regard to the final phonetic shape of the neuter and the locative clitics. Whereas in isolation these clitics surface either as vowels (in absolute initial position before a consonant) or glides, when in combination there is only one option, as shown in (16):[8]

(16)  a.  [**te**      **w**]      compres
          REFL.2S NEUT buy.2S
          *You buy this for yourself*

      b.  [**me**    **j**]    porta
          ACC.1S LOC bring.3S
          *S/he brings me there*

The fact that the locative and the neuter appear as glides in these combinations is a clear indicator that they must be underlyingly /j/ and /w/, because otherwise the insertion of the features corresponding to the theme vowel of the first clitic would be avoided. That is, in isolation, they surface as glides when followed or preceded by a vowel (e.g., *No* [**w**] *porto* 'I don't bring it'; [**j**] *arribo* 'I get there'), but this could be considered just a regular phonological process which is widespread crosslinguistically to avoid a hiatus. If these clitics were underlyingly vocalic, they should appear in

---

[8]See 6.3 for an analysis of the behavior of this clitics in isolation.

combinations as vowels because forms such as **\*[tu]** *compres* or **\*[mi]** *porta* are structurally optimal (in fact, these are the regular forms for Central Catalan). We must conclude, then, that input-output feature identity (IDENT-FEAT) overranks epenthesis of new phonological features, as shown in (17):[9]

(17)    *Input-output faithfulness*

| /m **V** # j **V** # p ɔ r t a/ | DEP-V | σSTR | ID-FEAT | DEP[-BK] | DEP-LK |
|---|---|---|---|---|---|
| a. ☞  [-back]<br>       \|<br>   m  e   j   p ɔ́ r t a | | | | 1 | 1 |
| b.    m     j   p ɔ́ r t a | | 1W | | L | L |
| c.    m    i   p ɔ́ r t a | | | 1W | L | L |
| d.    [-back]  [-back]<br>         \|        \|<br>   m  e   j  e   p ɔ́ r t a | | | | 2W | 2 |
| e.    [-back]<br>         \|<br>   m  e j   p ɔ́ r t a | 1W | | | 1 | 1 |

Candidate c. cannot be the winner because, as I just mentioned, it does not maintain the featural identity of the locative (IDENT-FEAT >> DEP[-BACK]), and candidate b., which does maintain it, is ruled out by σSTRUC. Candidate a. violates both DEP[-BACK] and DEP-LINK but it is the winner anyway because d. and e. must also be discarded due to the same reasons exposed above for the rest of the combinations (i.e., epenthesis of place features is not needed in two different theme positions and full epenthesis is dispreferred in Pallarese).

There is only one counterexample that seems to go against this hypothesis by which the neuter and the locative are underlyingly glides. The combination of *ne* + *hi* should yield, according to the ideas exposed above, the output \*[nej]. The prediction is not borne out and we find, on the contrary, the cluster [ni], where the

---

[9]The constraint IDENT-FEAT penalizes featurally unfaithful input-output mappings. Glides, as opposed to vowels, are [–syll], and thus the outputs [i]/[u] for the locative/neuter, which are [+syll], will incur in violations of IDENT-FEAT.

locative appears in its vocalic form, as shown in (18):

(18)   [**n**   **i**]   vam   menjar
       PART LOC go.1PL eat.INF
       *We ate there some of that*

The reasons for this unexpected behavior are not clear, but we could hypothesize that this structure is lexicalized as a whole. The existential construction *Hi ha* 'There is' appears very often in combination with the partitive *en* (i.e., *N'hi ha* 'There is (some of them)'). The fact that the phonetic output for it is [njá], without a vowel -*e*, and its high-frequency appearance, might play a role in the non-use of *[nej] in other contexts outside the existential construction. That is, some kind of frequency effect could ban the use of a form *[nej] even if it is the optimal candidate, given the constraint ranking in (17), because it differs from the form of the output cluster *N'hi ha*, where the combination of these two clitics is more commonly found. Nevertheless, this is only conjecture and deserves further investigation.[10]

## 6.2   1st and 2nd person plural

1st and 2nd person plural clitics have some peculiarities that need to be analyzed in detail. In 2.2 we saw how both pronouns behave similarly but, nevertheless, their forms differ in certain phonological environments.

The 1st person plural accusative and dative clitic appears always as *mos* in all circumstances:

(19)   a.   Les      que [**moz**]   van   acabar   de xafar
            ACC.3PL that ACC.1PL go.3PL finish.INF of crush.INF
            'Those things that finally crushed us'

       b.   No [**moz**]   interessa   la        política
            NEG DAT.1PL interest.3S the.FEM.S politics.FEM.S
            'Politics are of no interest for us'

As shown in (19), phonological conditioning plays no role in the final output of the 1st person plural clitic. Whether it appears before or after a vowel, or before or after a consonant, the output is always *mos*.

For the 2nd person plural clitic we find, instead, two forms, *vos* ([bos]) and *us*

---

[10]The morphological templates defended in Bonet (1991) for clitic combinations divide clitics into two fields, field A and field B. Field A includes all clitics specified with the feature [PERSON] (including *se*), while field B is occupied by the other clitics (3rd person, genitive, locative). As Bonet (p.c.) points out, the fact that the clitic *hi* and *en* are both integrated into what she called field B may favor the lexicalization hypothesis suggested here.

([ws]), where *vos* appears in pre- and postconsonantal environments:

(20)    a.    Estan  vigilant-  [**bos**]
              be.3PL guard.GER ACC.2PL
              'They are watching you'

        b.    [**bos**]  quedareu    aquí
              INH.2PL stay.FUT.2PL here
              'You'll stay here'

When preceded by a vowel (and followed by either a consonant or a vowel), however, the allomorph *us* ([ws]) is chosen:

(21)    No  [**wz**]    allunyeu!
          NEG REFL.2PL get.away.2PL
          'Don't get away!'

Given the distribution of 1st and 2nd plural clitics, we need to address the following related questions, which are relevant both phonologically and morphologically:

1. Both clitics contain a segment /z/ that is the expression of plural in the whole nominal system but, nevertheless, their final outputs cannot be considered fully compositional, i.e., they do not consist of 1st and 2nd person morphemes (as they appear in the singular) plus the plural marker. Thus, how should we segment them?

2. What is the underlying representation of these morphemes, and how can we account for their similarities and differences (e.g., can we derive the phonetic outputs of the 2nd plural clitic from one single underlying form)?

These two related questions are important for the correct understanding of the morphological and phonological characteristics that these clitics display. Therefore, the first task is, as I said, to determine their morphological decomposition. The 1st person plural clitic, for example, allows, in principle, the following segmentations:

(22)    a.    /mo+z/
        b.    /m+o+z/
        c.    /moz/

The 1st person singular clitic is /m/ and in (22) the segment /m/ appears again in its plural counterpart. Thus, there seems to be correspondence between the singular and the plural in the phonological spell-out of the first person. What is more confusing, though, is the appearance of a vowel /o/, whose morphological affiliation does not seem very clear. We have seen in chapter 5 that *-o* is the exponent for masculine

gender, which is used as a repair strategy in certain phonological environments in the masculine pronominal clitic and definite article *l/lo*. If we assume that /m/ is the exponent for 1st person and /z/ the exponent for plural, we could argue that the vowel *-o* in the 1st person plural clitic is also a case of morphological epenthesis. This hypothesis, however, proves to be wrong on two counts:

- Gender is not a featural component of 1st and 2nd person clitics, only of the 3rd person

- There is no phonological conditioning in the appearance of the vowel given that it is present in all contexts (pre- or postvocalically, pre- or postconsonantally)

In Greenberg (1963) and Harley & Ritter (2002) it is shown how, crosslinguistically, it is common for 1st and 2nd person pronouns not to express gender, as opposed to 3rd person pronouns. If gender is absent in the featural composition of the 1st person plural clitic it is obvious, then, that there cannot be any phonological exponent that matches it. I propose in this thesis that the phonological features associated with gender are floating (*-o* for the masculine, *-a* for the feminine) and only linked to skeletal positions under certain conditions, but the total absence of [±fem] in the morphosyntactic input leaves no space for this option. This is, as already noted, the same situation that we find in 3rd person plural accusative clitics and definite articles, which undergo Impoverishment of gender and, thus, do not have access to the phonological features provided by [±fem]. Moreover, the clitic *mos* is used for both masculine and feminine referents, and *-o* is the exponent for masculine only, which seems to discard as well the morphological epenthesis hypothesis.

In some varieties of Central Catalan and in Standard Catalan, the situation is different because, in fact, there is phonological conditioning. The 1st person clitic can have three different forms, shown in (23), depending on the phonological environment and whether it is proclitic or enclitic:

(23)    *ens, 'ns, -nos*

If the clitic is in absolute initial position, it appears as *ens* whether it is followed by a consonant or a vowel, as shown below:

(24)    a.    [**ənz**]    agrada
            1PL.DAT like.3S
            'We like (that)'

    b.    [**ənz**]    vestim
            1PL.REFL dress.1PL
            'We get dressed'

The schwa that appears at the beginning of the clitic is a product of epenthesis, something common in the clitic system of Central Catalan in initial position (it appears in all forms except for *us*: [əm], [ət], ...). The contact of the proclitic with the verb yields an illicit onset cluster, so word-initial epenthesis is the only possible solution.[11]

If there is a preceding vowel, instead, as in (25), it surfaces as *'ns* because there is no need for epenthesis:

(25)   a.   No   **[nz]**   agrada
            NEG  1PL.DAT    like.3S
            'We don't like (that)'

       b.   No   **[nz]**   vestim
            NEG  1PL.REFL   dress.1PL
            'We don't get dressed'

In enclitic position, however, we find the non-transparent form *nos* when the clitic is attached to the infinitive or the gerund (for some speakers/varieties of Central Catalan, not for all) as in (26):

(26)   Portar-      **[nus]**
       bring.INF    1PL.DAT
       'To bring us'

The contact with the infinitival suffix *-r* generates an ill-formed coda cluster – *[purtárns]– that is repaired with the use of *-o*, which seems, a priori, the same strategy used by Pallarese 3rd person singular clitic *l/lo*.[12] The problem, again, is that *-o* cannot be linked to gender because this is absent in the 1st and 2nd persons and it cannot be considered either a 'regular' epenthetic vowel due to its location and its quality (we would expect a schwa). Thus, it seems unclear as to why it appears in this context, but some kind of morphological conditioning seems to be at play.[13]

---

[11]For a detailed analysis of epenthetic processes in Catalan clitics see Bonet & Lloret (2002) and Bonet & Lloret (2005).

[12]The complex coda cluster *-rns* is sometimes tolerated in Catalan, e.g., [kárns] 'meat(s)'. Bonet & Lloret (2005) attribute the special behavior of clitics to the constraint ALIGN-RIGHT(SUB-$\sigma$) ("Align (Lex, R; M, N, R), the right edge of a lexical word (Lex) has to coincide with the right edge of some subsyllabic constituent, margin (M) or nucleus (N)")(p. 61), which is violated by complex subsyllabic constiutents. In the word *carns* the complex coda is part of the same lexical word, whereas in *[purtárns] the lexical word is *portar*, and thus the syllabification of the clitic violates ALIGN-RIGHT(SUB-$\sigma$).

[13]It is not clear either why the infinitival suffix *-r* is pronounced in these cases, as in isolation it never surfaces ([purtá]; final *r*-deletion is a widespread process in the phonology of Catalan). Mascaró (1986) assumes that in verb + clitic sequences the prosodic word includes the clitic and the context for *-r* deletion is thus not met anymore.

What I want to illustrate with these examples is that the behavior of Pallarese cannot be grouped together with that of Central Catalan. There are no specific variants for different phonological contexts in Pallarese Catalan (syllable structure does not need to be 'repaired') and, thus, the only possible segmentations are those of (22), where *-o* is always present in the underlying representation. The main problem, though, is determining whether /z/ is part of the same lexical item or whether it is an independent morph. The answer to this question is not uniform in the literature, and while Mascaró (1986) considers that /z/ is the exponent for plural, Bonet & Lloret (2005) argue, instead, for a monomorphemic clitic that includes the sibilant as part of its segmental inventory (i.e., (22-c)).

Even if (22-c) seems a possible option, the behavior of 1st and 2nd person plural clitics in combinations with *hi* and *ho* (analyzed in 6.2.1) favor an interpretation by which the 1st person pronoun must be segmented like (22-b), with the vowel *-o* occupying a theme position and a number projection filled by the plural morph /z/. This is the segmentation provided in (22-b), which departs from Mascaró (1986) by positing a theme vowel. That is, Mascaró assumes a structure like (22-a), where *-o* is part of the allomorph for 1st person plural, but he does not consider it a separate morpheme, which I have labelled theme vowel in this thesis. Therefore, the morphological structure for the clitic *mos* must be the one shown in (27):

(27)

```
            #
          /   \
        D       #
       / \      |
      D  Th     z
      |   |
      m   o
```
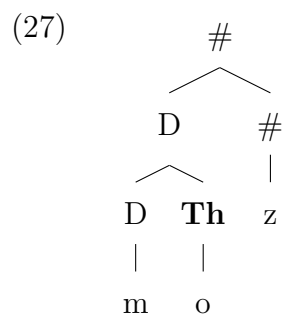
What this morphological structure shows is that, in fact, there is allomorphy for the 1st person morpheme depending on number. That is, there are two allomorphs, /m-V/ and /mo/, the first one being used in the singular and the second one in the plural (this is, thus, a case of morphologically conditioned allomorphy). The theme position is left unspecified in the singular, whereas in the plural there is a lexicalized theme vowel *-o* which is in no way linked to the general process of gender inflection that provides the place features for *-o* in the masculine and the place features for *-a* in the feminine. The vowel *-o* is here, instead, stored with the lexical entry, where it occupies the theme position.[14] The motivations for the choice of this structure,

---

[14]We will see in section 7.1 of chapter 7 how this idea of lexicalized vowels in the theme position is applied to all non-regular elements of the nominal system.

as I said, will be clearer in section 6.2.1, devoted to clitic combinations.

Until now we have analyzed the 1st person plural clitic, but the 2nd person plural clitic, which poses more problems for interpretation, has remained unexplored.[15] As it has been shown in (20) and (21), we find two different phonetic outputs for the clitic, and the question that needs to be answered is whether we need to resort to allomorphy or we can derive both forms from one single underlying representation. Bonet & Lloret (2005, p. 47) assume that, for some varieties of Central Catalan, there is only one underlying form for the 2nd person clitic, /wz/, which does not have internal structure. Glides cannot be syllabic in Catalan and, as a consequence, in absolute initial position or after a consonant the 2nd person plural clitic becomes vocalic, as in (28):

(28)   [**us**]     compro  un        gelat
       DAT.2PL buy.1PL a.MASC.S ice-cream.MASC.S
       'I buy you an ice cream'

This is the strategy used by Central Catalan to avoid an initial glide, whereas after a vowel the faithful output is maintained (*No* [ws] *compro un gelat* 'I don't buy you an ice cream'). If the underlying representation for the 2nd person plural clitic contained a glide in Pallarese as well, it is clear that the strategy to avoid it in absolute initial and postconsonantal contexts would be different. The form employed by Pallarese is [bos], which does not seem easily derivable from /wz/. However, with some intermediate steps, we could even do so.

It is commonly assumed that glides can imply a certain degree of frication (Ladefoged & Maddieson, 1996; Padgett, 2008). This means that, in the phonological contexts just mentioned, the feature [+labial] could trigger the appearance of the bilabial approximant, as shown below:

(29)   /wz/ > βuz

This would imply that the glide is split into a consonant and a vowel, thus creating a regular onset + nucleus (+ coda) syllable structure. However, approximants are not tolerated in absolute initial position in Catalan, and, therefore, we need one more step in the derivation that converts the approximant into a stop (by changing the continuancy value), as in (30):[16]

(30)   /wz/ > βuz > buz

---

[15]The morphological structure just advocated for the 1st person plural clitic is also valid for the 2nd person plural.

[16]Stop/approximant alternations are pervasive in the phonology of Catalan and the same general processes would apply here.

The clitic, though, is not *[bus] but [bos], with a mid-back vowel. What could trigger the change in the vowel quality? If, as we saw, the clitic for the 1st person plural is *mos*, with the same mid-back vowel *-o*, it could be argued that paradigmatic pressure has an influence in the final outcome of the 2nd person plural clitic. Both clitics are very similar and group together in many respects, and this could force a reinterpretation in the system by which they have to pattern together phonologically in certain contexts, thus changing /u/ to /o/. All these derivations, though, do not seem very straightforward and it is more plausible that both variants, /wz/ and /boz/ are simply allomorphs of the 2nd person plural (Mascaró 1986 assumes the same for Central Catalan, with two underlying representations {/u/,bu/}, although the distribution of the allomorphs is different —the form *vos* only appears after gerunds or infinitives in this dialect, e.g., *donant*-[bus] 'Giving (this) to you').

Opting for an allomorphic solution also poses some problems, though. As already stated, the difference in exponence between the 1st person singular and the 1st person plural (also extendable to the 2nd person), is a case of morphologically conditioned allomorphy. In addition, in the 2nd person plural we are facing a case of phonologically conditioned allomorphy, as there are two different outputs for the same set of morphosyntactic features which are dependent on the phonological context. The underlying representations are the following:[17].

(31)    a.    /b–o+z/
        b.    /w–V+z/

The form in (31-a) behaves like the 1st person plural clitic *mos*, i.e., it contains a vowel *-o* (not linked to gender) which occupies the theme position. (31-b), by contrast, maintains the theme position with the underspecified vowel found across the nominal system. As this allomorph appears only after a vowel, the glide can still be attached to the previous host (as a coda, also with the plural morph: *No'*[wz] *allunyeu* 'Don't go away'), and thus the V-slot cannot be phonetically interpreted because it lacks place features.

The main problem for this kind of allomorphy is how to determine the correct selection of a form for every specific context. Mascaró (1996) and Bonet et al. (2007), among others, analyze phonologically conditioned allomorph selection in terms of optimization. That is, certain allomorphs are preferred over others because they improve syllable structure, for example. Phonology, thus, has an impact on the

---

[17]I follow Viaplana (1980), Bonet & Lloret (2005) and Massanell (2011) in considering that the glide is part of the underlying representation in (31-b), although in my analysis there is also a consonantal allomorphic variant, i.e., (31-a) (i.e, I do not consider that all phonetics forms can be derived from one single underlying form /w+z/ or /wz/)

surface form of a given morpheme.[18]  Although the shape of the 2nd person plural clitic in Pallarese is phonologically conditioned, with *us* appearing only after vowels and *vos* elsewhere, it seems difficult to support an optimizing analysis. Take, for instance, the sentence *No'*[wz] *allunyeu* 'Don't go away', which has the following input (the internal structure of the clitic is omitted):

(32)    /no#{wz,boz}#aʎuɲɛw/

In (32) the two allomorphs are in the input, and therefore the constraint ranking should favor one of the two forms. Why should /wz/ be preferrable to /boz/ only after vowels on phonological grounds? To my understanding, selecting /wz/ does not improve, structurally speaking, the phonological sequence. In fact, /bos/ contains an onset and should thus be favored by markedness from a syllabic point of view, since the attachment of /wz/ to the previous vocalic host generates a complex coda in the syllable ([nowz]). The form /boz/ seems the most adequate allomorph in all contexts, and there is no straightforward explanation for the use of /wz/ in postvocalic environments.

Paster (2006) and Bye (2007) reject the idea of optimization in phonologically conditioned allomorph selection and propose instead a subcategorization-based approach for this kind of allomorphy in which "phonological conditions on affixation are modeled by incorporating phonological aspects of stems directly into the selectional requirements of affixes" (Paster, 2006, p. 20). Thus, in such an approach the lexical entries for the Pallarese 2nd person clitic look as in (33):

(33)    a.    2PL ↔ /wz/ /V__
        b.    2PL ↔ /boz/ (elsewhere)

This implies, *de facto*, that all phonological generalizations are lost because each phonological condition is lexically listed for a particular form, which makes the situation completely arbitrary. Therefore, the reverse specifications could be found for the 2nd person plural in (33) (i.e., *vos* after a vowel and *us* elsewhere).

The total denial of phonological optimization posited by subcategorization models seems too strong (too many allomorphic patterns point in this direction to just consider it a fortuitous phenomenon). Non-optimizing allomorph choice is better captured in subcategorization frameworks and other authors (e.g., Booij 1998 and Nevins 2011) consider that both options can coexist. In some cases, output conditions are imposed on morphology (i.e., phonology determines the choice of an allomorph), but in others, morphology itself imposes certain structural requirements

---

[18]This is the P >> M approach, see 1.1.3 for more details.

on complex word formation. As there seems to be no phonological motivation for the use of /wz/ after vowels, I assume the Vocabulary Items in (33) for Pallarese in which phonological conditionings are specified in the context of insertion. As just mentioned, this does not imply that phonological optimization does not play any role in the grammar.

### 6.2.1   Combinations

As we have already seen throughout this dissertation, the Person Case Constraint limits the number of possible clitic combinations in Catalan. In Pallarese, we find three kinds of clitic clusters which involve the 1st and the 2nd person plural:

1. *se* + 1st/2nd person plural.

2. 1st/2nd person plural + 3rd person accusative. The 3rd person accusative can appear as both masculine and feminine, singular or plural (recall that there is Impoverishment in the plural and thus no gender distinction).

3. 1st/2nd person plural + genitive/locative/neuter (i.e., *ne, hi, ho*)

Within the first group, exemplified in (34), we find the forms [mos] and [ws]:

(34)    *Reflexive + 1st/2nd person plural*

    a.   [**se   moz**]   arrupís l'       estómac
         REFL DAT.1PL turn.3S DEF.MASC.S stomach.MASC.S
         'Our stomachs get turned'

    b.   [**se   wz**]   arrupís l'       estómac
         REFL DAT.2PL turn.3S DEF.MASC.S stomach.MASC.S
         'Your stomachs get turned'

The outcome of these clusters is quite transparent. The clitic *se*, which is underlyingly /s/, undergoes epenthesis of vocalic features in the theme position (the same as happened in isolation in certain conexts) due to the contact with the following clitic. This epenthesized vowel, in turn, allows the appearance of the 2nd person plural clitic in its postvocalic form [wz], which is preferred over *vos* in this context. The problem that faces this analysis, though, is the following: if these features are only epenthesized due to the phonological context (the combination of the two clitics creates an ill-formed onset cluster), why do we not find outcomes like those of (35-a), where the complementizer *que* provides a host vowel for the reflexive?

(35)    a.   *que [**z moz**] arrupís l'estómac

    b.    que [**se moz**] arrupís l'estómac

Instead, the right sentence is (35-b), with epenthesized features in the theme position again, even if the phonological context seems to advocate the non-exisiting combination in (35-a) (the clitic *se* could be attached to the vowel of the complementizer). The answer to this puzzle is not simple, and it could be linked to the notion of phase (Chomsky, 1999). If, as Bonet (2013) notes, we assume Chomsky's notion of phase, the complement of the phase head, but not the head itself, is sent to the interfaces for interpretation, which correctly predicts (35-b), because the clitic cluster is phonologically spelled out before the complementizer, and thus the postvocalic context provided by it comes too late.[19] Newell (2008) argues instead that the head of a phase is interpreted at PF with its complement, which does not hold for the example just mentioned because it implies that the clitic would be in the same spell-out domain as the complementizer and could thus be attached to it. The picture is more complicated than it appears here, though, because even if (35) favors an interpretation by which the CP head and its complement are phonologically spelled out separately, other cases involving lexical insertion in clitics point in the opposite direction. We find many examples of these in Pallarese, as well in the 3rd person singular masculine pronominal clitic or definite article analyzed in chapter 5. The difference in the exponence with regard to (35) is shown below:

(36)    a.    [**lo**]    pa    és    bo
                DEF.M.S bread.MASC.S be.3S good.MASC.S
                'The bread is good'

        b.    que [**l**]        pa    és    bo
                that DEF.MASC.S bread.MASC.S be.3S good.MASC.S
                'That the bread is good'

        c.    que [**z**]    van    abraçar
                that REFL go.3PL hug.INF
                'That they hugged'

If the process of spell-out were the same in all contexts, we would expect *\*que lo pa es bo*, as the phonological context provided by the complementizer should only be available after the clitic had been already spelled out (this is what (35-b) suggests).[20] The same happens with the reflexive in (36-c), which can be encliticized

---

[19]One could argue that open syllables are preferred crosslinguistically and thus (35-b) is preferred over (35-a). We will see below that this does not hold true in other cases.

[20]A sentence like *que lo pa és bo* can actually be found in Pallarese, but it depends on pauses and other factors that influence casual speech. When there is a previous vocalic host, though, the form without the gender exponent is, by far, the most preferred.

to the complementizer.[21] It could be argued, maybe, that the fact that in (35) we have a clitic combination and in (36) the clitic appears in isolation could have an effect in the spell-out process. In fact, in combinations the clitics are always grouped together, as in (37):

(37)  que [**se  t**]'  escapen
      that REFL DAT.2S escape.3PL
      'That they go away from you'

A sentence like *\*que* [s t]' *escapen* —totally acceptable from a phonological point of view— is never found, so it seems there is an asymmetry between isolated clitics and combinations. It was shown in chapter 5, however, that 3rd person accusative + dative combinations behave like (36) (i.e., like the clitic in isolation), as shown below:

(38)  a.  Les  tomates,  diu  que' [**lzi**]  portes
         the.FEM.PL tomato.FEM.PL say.3S that ACC.3PL+DAT.3S bring.2S
         lo  dimarts
         the.MASC.S Tuesday.MASC.S
         'S/he says that you bring her/him the tomatoes on Tuesday'
    b.  Amb ell, [**lezi**]  portes  lo  dimarts,
         with  he  ACC.3PL+DAT.3S bring.2S the.MASC.S Tuesday.MASC.S
         les  tomates
         the.FEM.PL tomato.FEM.PL
         'You bring him the tomatoes on Tuesday'

Why do (35) and (36) & (38) pattern in a different way? Maybe the answer lies in the nature of the clitic cluster. 3rd person dative + accusative combinations are opaque in Pallarese Catalan and generate outputs that look like the 3rd person dative in isolation (i.e., *li* or *les hi*), which seems to have an effect in the spell-out of the morphosyntactic combination.[22] That is, the exponence of this bundle of pronominal features is a single clitic, and as such, it patterns phonologically with clitics in isolation –like (36) instead of (35)–. In fact, the asymmetries shown above seem to indicate that clitic clusters constitute a somehow independent spell-out group (Bonet, p.c.). If we assumed a cyclic model of phonology, certain operations would apply first to the clitic cluster, which defines a phonological domain, and then,

---

[21]Again, there is a lot of variation, and it is not uncommon to find the full form *se* in these contexts, i.e., *que* [se] *van abraçar*.

[22]Recall, however, that, at least in the plural, there is a subtle difference. While *les* is the preferred option in isolation, in combinations we mostly find *les hi*. Nevertheless, *les hi* is a possible option as a plural dative for some speakers, which means that it also has to be treated as a single clitic.

in the next cycle, other post-lexical rules would apply, but the clitic combination would have already been spelled out and would thus be blind to the phonological context provided by the complementizer in (35-a). The differences in the spell-out domains can be better understood in (39):

(39)    a.   [host + clitic]

          b.   [host + [clitic + clitic]]

This hypothesis by which clitic clusters do not pattern as isolated clitics seems to be borne out by the behavior of 1st/2nd person plural + 3rd person accusative (singular or plural) clitic combinations, which I reproduce again below (from chapter 2):

(40)    *1st/2nd person plural + 3rd person singular feminine*

      a.   Doneu-   **[mo   la]**,   la   tomata!
          give.2S.IMPR DAT.1PL ACC.3S.FEM the.FEM.S tomato.FEM.S
          'Give us the tomato!'

      b.   Mengeu-   **[βo   la]**,   la   tomata!
          eat.2S.IMPR INH.2PL ACC.3S.FEM the.FEM.S tomato.FEM.S
          'Eat the tomato!'

      c.   La   tomata   **[βo   l]**'   amaniu
          the.FEM.S tomato.FEM.S INH.2PL ACC.3S.FEM dress.2PL
          'You dress the tomato'

(41)    *1st/2nd person plural + 3rd person singular masculine*

      a.   Aguell   enciam,   porta-   **[mo   l]**!
          that.MASC.S lettuce.MASC.S bring.2A.IMPR DAT.1PL ACC.3S.MASC
          'Bring us that lettuce!'

      b.   Mengeu-   **[βo   l]**,   lo   lluç!
          eat.2S.IMPR INH.2PL ACC.3S.MASC the.MASC.S hake.MASC.S
          'Eat the hake!'

(42)    *1st/2nd person plural + 3rd person plural*

      a.   **[mo   ls]**   comprem
          DAT.1PL ACC.2PL.MASC buy.1PL
          'We buy ourselves these things'

      b.   Mengeu-   **[βo   ls]**!
          eat.2PL.IMPR INH.2PL ACC.3PL.FEM
          'Eat them!'

(43)    *1st/2nd person plural + 3rd person plural (-s variant)*

      a.   Les   tomates   no   **[mo   z]**   vol
          the.FEM.PL tomato.FEM.PL NEG DAT.1PL ACC.3PL.FEM want.3S

comprar dingú
buy.INF  nobody
'Nobody wants to buy us the tomatoes'

b.   Les          lluços,       **[bo s]**  porten      al
the.MASC.PL hake.MASC.PL DAT.2PL ACC.3PL.MASC bring.3PL
restaurant
to+the.MASC.S restaurant.MASC.S
'They bring you the hakes to the restaurant'

1st and 2nd person plural clitics are spelled out as *mos* and *vos*, with a sibilant, in isolation. We can see in the examples above, though, that they lose this final segment when attached to the 3rd person accusative clitic, in both their singular and plural forms. Only in (43) the sibilant appears on the surface, but it is the expression of the 3rd person accusative plural clitic, or at least the fusion of the sibilants of both the dative clitic and the accusative clitic (as shown in previous chapters, the accusative plural can appear as *-s* in certain contexts, as in *No* [**s**] *comprem* 'we don't buy these things').

The same process of *-s* deletion is found in the combinations which involve the clitic *ne*, as shown in (44):

(44)   a.   Anà- **[βo    n]**  d' aquí
go.INF INH.2PL PART of here
'To leave (you) a place'

b.   **[mo    n]**  comprem dos
DAT.1PL PART buy.1PL  two
'We buy two of those for ourselves'

It seems, thus, that the contact between a lateral or a nasal and a sibilant (/mo**z**/-/bo**z**/+/l(+z)/-/**n**/) causes the disappearance of the sibilant. This only happens, though, when two clitics are attached to each other, not when the 1st/2nd person plural clitic is attached, in isolation, to a word that begins with a lateral or a nasal, as in (45), which demonstrates that the realization of the clitic is independent from the word it is attached to:

(45)   **[moz]**   lamentem
REFL.1PL complain.1PL
'We complain'

The asymmetries between (40), (41) and (42), on the one hand, and (45), on the other hand, show that the clitic cluster constitutes an independent spell-out domain where certain phonological processes can apply. Outside of it, by contrast, these processes cannot apply (cf. (45)). Again, this phenomenon can be better explained

by cyclic models of phonology (e.g., Stratal OT, Bermúdez-Otero 1999; Kiparsky 2000; see Vigário 1999 for an analysis of cliticization in European Portuguese as a postlexical operation).

Finally, the combinations of 1st and 2nd person + locative and neuter deserve some specific comments. It was determined in section 6.1.1 that *hi* and *ho* are underlyingly glides which, in combinations, cause the insertion of epenthetic features in the theme position of the clitic they are grouped with (which appears in first place) in order to preserve their featural identity (i.e., in order to remain glides). Thus, in combination with *me*, *te* or *se* we find the forms [mej]/[mew], [tej]/[tew], [sej]/[sew], where the theme position of the pronouns /m-V/, /t-V/ and /s-V/ is used as a 'rescue' strategy to maintain the glides in the output (which cannot be syllabic themselves in Catalan, as opposed to vowels). In clusters that involve 1st and 2nd person plural clitics + locative/neuter, however, the expected glides are replaced by vowels, yielding forms such as [mozi]/[mozu] and [bozi]/[bozu], as in (46):

(46)   a.   [**moz**   **i**]   porta   ara
         ACC.1PL LOC bring.3S now
         'S/he takes us there'

       b.   [**boz**   **u**]   diu
         DAT.2PL NEUT say.3S
         'S/he tells us this'

Why do these clitics behave differently from the other set of pronouns? The answer to this question, which may not seem obvious at first sight, lies, once more, on the morphological structure of the plural clitics.

Recall, once again, that every syntactic functional head is assumed to project a theme position, which we determined to be occupied by the lexicalized vowel *-o* for the 1st and 2nd person plural clitics, as shown in (47) with the clitic *mos*:

(47)
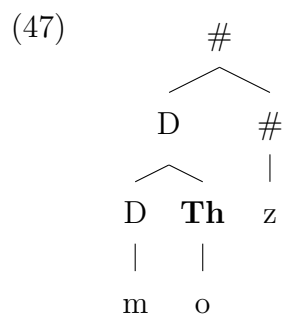


Number, as depicted in (47), does not project a theme position. This implies that there is no morphological slot available to insert vocalic place features in case of a syllabification problem, and the only alternative is to epenthesize a full vowel,

which proved to be highly costly in Pallarese. In fact, the absence of this position is the reason why we find outputs such as [mozi] and [bozi]. Neuter and locative clitics are underlyingly glides, but their attachment to *mos* and *vos* would yield unsyllabifiable sequences such as *[mozj] and *[bozj].[23] The insertion of *-e* (e.g., *[moz**e**j]) could maintain a faithful input-output mapping for the glide, but in the constraint ranking DEP-V dominates IDENT-FEAT (a constraint that argues for the preservation of input features in the output), which implies that the vocalic candidate, candidate a., is the winner, as shown in (48):[24]

(48)     **[mozi]** porta

| /m o + z # j **V** # p ɔ r t a/ | DEP-V | σSTR | ID-FEAT | DEP[-BK] | DEP-LK |
|---|---|---|---|---|---|
| a. ☞ m o + z   i   p ɔ́ r t a | | | 1 | | |
| b.   m o + z   j   p ɔ́ r t a | | 1W | L | | |
| c.   m o + z   j e   p ɔ́ r t a     [-back] | | 1W | L | 1 | 1 |
| d.   m o + z   e j   p ɔ́ r t a     [-back] | 1W | | L | 1 | 1 |

There are several issues that need to be accounted for in this tableau, which highlight the complex intertwining between syllabic and morphological structure. First, as already noted, there is no theme position available for the 1st person plural because number does not project it, and the theme position of the D head is already spelled out by the lexicalized vowel *-o*. This is the reason we only find an underspecified vowel for the locative, /j–V/, in the input. In fact, (48) shows that the morphological segmentation that has been assumed for *mos* and *vos*, where /z/ is

---

[23]It has been claimed in the literature that glides are always positional variants of vowels, which are their only underlying source; that is, underlying glides do not exist (Rosenthall 1997, but see Levi 2008 for an opposing view). The behavior of the neuter and locative in combinations (see section 6.1.1) seem to point in this direction, though, and I am forced to conclude that a glide must be present in the underlying representation of these clitics. Underlying glides have also been posited for the 2nd person plural suffix /w/ in Catalan verbal morphology (Mascaró, 1986) and the 2nd person plural clitic in Viaplana (1980), Bonet & Lloret (2005) and Massanell (2011) (as mentioned above), but this issue is controversial and deserves further investigation.

[24]I am using σSTR in a 'lax' way in the following tableau, as it is understood here as a cover constraint for different phonotactic restrictions that control syllable structure in Pallarese Catalan.

an indepenent plural morph, and not just a segment of the features that correspond to the whole 1st and 2nd person plural clitcs, is correct because it predicts that the output for the combination will be [mozi] (or [bozi]). If, instead, we assumed the clitic to be monomorphemic, we would have the input /moz–V/+/j–V/ for the cluster, and therefore the winning candidate would be *[mozej], with epenthesis of place features that would end up linked to the theme position of the 1st person plural clitic.

In (48), the featural composition of the locative, that corresponds to a glide, should be preserved, but the syllabic contacts it creates with the first clitic of the cluster bans this option. Candidate b., for instance, is ruled out because /j/ cannot be syllabic (the same happens in combinations with *me*, *se* and *te*, as shown in (17)), and thus violates $\sigma$STR. In candidate c., by contrast, the glide can be considered both an onset itself and the second element of a complex onset, depending on the syllabification we assume for the cluster ([moz.je] or [mo.zje]; I did not include two candidates because both can be ruled out by the cover constraint $\sigma$STR, which encodes here different restrictions that control Catalan syllabic contacts). The syllable contact law (Vennemann, 1988) disfavors a candidate [moz.je], with rising sonority in heterosyllabic sequences (see Pons-Moll 2011 for a thorough analysis of syllabic consonantal contexts in Catalan and other Romance languages). If we assume, on the contrary, that the cluster is syllabified as [mo.zje], we still have a problem. Complex onsets are dispreferred crosslinguistically, and this also holds true for the present sequence of fricative+glide, where a candidate with the glide as part of the onset can be ruled out by the action of a constraint such as *ONS/GLIDE (Colina, 2007; Bradley, 2014) which goes against glides surfacing as onsets/complex onsets (ONS/GLIDE is grouped within $\sigma$STR, and therefore candidate c. violates this constraint).

Therefore, it seems that Pallarese Catalan penalizes glides as onsets but not as codas, although we see in (48) that candidate d. is also ruled out, even though [j] appears in coda position. The reason is, again, the presence of the high-ranked constraint DEP-V, which militates against full epenthesis. If the theme position of the pronoun *mos* were available, a candidate [mozej], with epenthesis of place features only, would be the winner, as IDENT-FEAT dominates DEP-[BACK], but this option is not available as the number layer does not project a theme position.

The behavior of these clitic clusters, thus, acts according to the predictions made by the present analysis of the nominal system of Pallarese and supports the epenthesis-as-last-resort strategy argued for in this dissertation. To my understanding, it is not a coincidence that, whenever there is a theme position available, the

insertion of unmarked vocalic features is used to preserve the featural identity of the locative/neuter. When there is no such position, the alternative option is to epenthesize a whole vowel (skeletal structure included), a process which proves to be highly costly and, in fact, avoided even when it causes a featural mismatch in locative/neuter input-ouput mappings. The tableau in (48) shows how close phonology and morphology work together but, at the same time, it also shows how they maintain their independence, as phonology can only make use of morphology in very specific conditions.

## 6.3 Neuter and Locative

Neuter and locative clitics have the same morphological structure as the other pronouns analyzed in this chapter, i.e., they project a theme position, as we have already seen. The clitics *ho* and *hi* consist of a glide (/w/ and /j/, respectively) and have been already analyzed in combinations with other clitics. Only a few remarks will be made in this section to fully understand their morphological and phonological behavior.

Even if the neuter and the locative are underlyingly glides, they surface as vowels in certain environments. In isolation, we find [u] and [i] in absolute initial position before a consonant, as in (49):

(49)   a.   **[i]**   treballàvom
           LOC work.IMPF.1PL
           *We worked there*

       b.   **[u]**   pensa
           NEUT think.3S
           *S/he thinks so*

To preserve the featural identity of the clitics, the forms [je] and [we] should be favored, but we find their vocalic counterparts instead. The theme position is available and could be used to epenthesize the vocalic features of the mid-front vowel -*e*, but the constraint *ONS/GLIDE dominates IDENT-FEAT, and thus [je] and [we] are discarded.

(50)    [**u**] pensa

| /w **V** # p ɛ n s a/ | Dep-V | σStr(*O/Gl) | Id-Feat | Dep[-bk] | Dep-Lk |
|---|---|---|---|---|---|
| a. ☞  u    p ɛ́ n s a | | | 1 | | |
| b.  w    p ɛ́ n s a | | 1W | L | | |
| c. [-back]<br>w e    p ɛ́ n s a | | 1W | L | 1 | 1 |
| d. [-back]<br>e w    p ɛ́ n s a | 1W | | L | 1 | 1 |

The tableau in (50) is parallel to that of (48), as the restrictions that force a vocalic output in the combination *mos + hi* are equally at play here. Candidate d. is ruled out due to the dispreference for the epenthesis of a full vowel, whereas the Sonority Hierarchy prevents candidate b. from being the winner. Although candidate a. violates the faithfulness constraint IDENT-FEAT, it is preferred over candidate c. due to the tendency to avoid glides as onsets (*ONS/GLIDE).

However, we do sometimes find the locative and the neuter as onsets, but only when a hiatus would generate a worse syllable contact. This is what we find in examples like [w] *agafa* 'S/he takes it' and [j] *arriba* 'S/he gets there', where the glide as an onset is preferred over a hiatus (*[u.a]*gafa*, *[i.a]*rriba*), as shown in (51):

(51)     [**w**] agafa

| /w **V** # a g a f a/ | Dep-V | *V.V | *O/Gl | Id-Feat | Dep[-bk] | Dep-Lk |
|---|---|---|---|---|---|---|
| a. ☞ w    a ɣ á f a | | | 1 | | | |
| b.    u   a ɣ á f a | | 1W | L | 1 | | |
| [-back]<br>    \|<br>c.  w e   a ɣ á f a | | 1W | 1 | | 1 | 1 |
| [-back]<br>    \|<br>d.  e w   a ɣ á f a | 1W | | 1 | | 1 | 1 |

The constraint *V.V ("two consecutive heteromorphemic vowels are not allowed") dominates *Onset/Glide and, therefore, candidate b. is ruled out.[25] As there is no better candidate to avoid the appearance of [w] as an onset, candidate a. is the winner, because candidate c., with epenthesis of vocalic features, also contains a glide and, in addition, it generates a hiatus (*[we.a.ɣá.fa]). Candidate d., even though it maintains the featural identity of the clitic, is discarded due to Dep-V.

Most combinations that involve the neuter and locative clitics have already been anlayzed in previous sections. Only the following combinations need to be accounted for:

- ho + hi

- ne + ho

Both combinations are rare in all the Catalan-speaking domain. That is, they are not impossible but are only elicited in very specific contexts, as opposed to all the other combinations that have appeared throghout this chapter which are very frequent.

As pointed out by Bonet (2002), the combination *ho + hi*, which can arise when an argument and an adjunct of the verb are pronominalized, is usually not attested

---

[25]In th feminine accusative clitic *la*, *V.V also accounts for the non-presence of -[a] in prevocalic contexts in absolute initial position, e.g., *l[a a]*gafo* 'I take it'. See section 5.1.1 for more details.

in Catalan except for the Ribagorçan dialect, spoken in Northwestern Catalonia and Northeastern Aragon (*Allí **ho hi** tinive de qualsevol manera* 'S/he had that thing there carelessly' —Suïls (1993)). Pallarese and Ribagorçan are in contact, but I could not elicit any sponataneous production of this clitic cluster.

As for the other combination, *ne + ho*, it is also usually avoided in Catalan. Bonet (2002) provides examples of the sequence *n'ho* only when the clitic *ne* is inherent (*endur-se-**n'ho*** 'take something away'), never when it is an argument. The solutions attested to avoid this combination vary between dialects, and consist of the deletion of *en* (*en + ho = ho*), the substitution of *ho* for *l* (*en + ho = l'en*) or the substitution of both clitics, *ho* for *l* and *en* for *hi* (*en + ho = l'hi*). Again, this combination is not easily elicited and I did not find the sequence *ne + ho* in any of the recordings.

Be that as it may, the problems associated with these clitic combinations are basically morphological, not phonological, and are therefore left for further investigation.

# SEVEN

# CONSEQUENCES OF THE PROPOSAL

## 7.1 Autosegmental inflection in nouns and adjectives

### 7.1.1 Noun classes

It has been assumed throughout this dissertation that functional categories project a theme position postsyntactically. Recall, once again, that this theme position is realized by an underspecified vowel linked to a timing unit, which ends up attracting the floating phonological features that correspond to gender values:

(1)  a.  Underspecified vowel in the theme position:
         Th ↔ V (*à la* Roca 2010)
     b.  Association of gender values with phonological features:

         | [+fem] | ↔ | [+low] | (= /a/) |
         |--------|---|--------|---------|
         | [–fem] | ↔ | [+labial, –high] | (= /o/) |

I claimed that this process of phonological spell-out could explain the regular cases of the nominal system, i.e., the feminine nouns with an *-a* ending, like *casa* 'house', and masculine nouns with a *-Ø* ending, like *nen* 'kid' (the ranking DEPLINK >> PARSE[+LABIAL] prevents /o/ from being linked to the theme position in the masculine, whereas in the feminine PARSE[+LOW] >> DEPLINK ensures that the exponent is overt). The remaining nouns, which fall out of the default pattern and bear different inflectional endings, are all lexicalized. However, I have not elaborated further on this issue, which is crucial for a complete and exhaustive analysis of the Catalan nominal system.

First of all, we need to determine the inflectional endings of Pallarese Catalan. For the sake of clarity, I reproduce them again below:

| | | Masculine | | Feminine | |
|---|---|---|---|---|---|
| Inflectional marker | [a] | *poet*[a] | 'poet' | *cas*[a] | 'house' |
| Inflectional marker | [Ø] | *cor*[Ø] | 'heart' | *sal*[Ø] | 'salt' |
| Inflectional marker | [o] | *llor*[o] | 'parrot' | *mot*[o] | 'motorcycle' |
| Inflectional marker | [e] | *par*[e] | 'father' | *mar*[e] | 'mother' |

Table 7.1: *Pallarese inflectional endings*

The endings presented in table 7.1 are the most frequent and widespread suffixes in Pallarese (with differences, though, because -[a] and -[o] are very rare in the masculine and feminine, respectively), but the following markers, although peripheral in the system, can also be found (as in the whole Catalan-speaking domain): *-u*, *-i*, *-s*, *-is*, *-us*, *-es*, *-os*, as exemplified in table 7.2:[1]

| | *u* | *i* | *s* | *is* | *us* | *es* | *os* |
|---|---|---|---|---|---|---|---|
| Masc. | *dé*[w] | *bigot*[i] | *temp*[s] | *brind*[is] | *cact*[us] | *Sòcrat*[es] | *cosm*[os] |
| Fem. | *trib*[u] | *cris*[i] | *pol*[s] | *glot*[is] | *Ven*[us] | *càri*[es] | – |

Table 7.2: *Non-canonical endings in Pallarese Catalan*

These marginal nouns belong to the athematic class (in Bermúdez-Otero 2006b —and subsequent work— terms), i.e., they do not bear the regular theme vowels, and they can appear with both masculine and feminine nouns, except for *-os*, which is only found in the masculine. Given the distribution of these inflectional markers, a fundamental question needs to be answered: what is the relationship between syntactic gender and the different vowel endings of Pallarese Catalan?

As already mentioned, the hypothesis I am pursuing is that [+fem] is realized by /a/ and [–fem] by /o/. Thus, there is a direct correlation between syntactic gender and inflectional endings in the regular default cases. This does not mean, though, that all other endings are formal class markers (*à la* Harris 1985). All these different inflectional markers are considered, instead, either a) theme vowels stored in the theme position in lexical entries or b) part of the root. The differences of each group are analyzed in sections 7.1.1.1 and 7.1.1.2 depending on their morphological and

---

[1]The *-u* marker usually appears as -[w] in masculine nouns. Lloret (2013), who provides a complete review of the inflectional system of Catalan, distinguishes between *-w* and *-u* (which appears in loanwords such as *Sudok*[u]).

phonological behavior.[2]

### 7.1.1.1  Default nouns

Default nouns, which bear -*a* in the feminine and -*Ø* in the masculine, do not pose any problem for the proposal because they behave as expected. As already mentioned throughout this dissertation, I adopt Oltra-Massuet (1999)'s morphological structure for Catalan nouns, where the theme position is filled with an underspecified vowel:

(2)

```
                    n
            _____|_____
           /                       \
       √ROOT                        n
                            _____|_____
                           /                 \
                      n_[±fem]             Theme
                                              |
                                              V
```
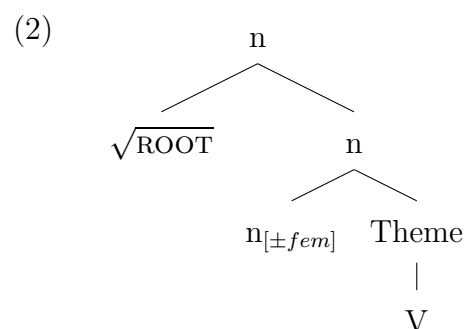
This structure correctly predicts the realization of the phonological features associated with either feminine ([+low]) or masculine ([+labial]) in the V-slot, given the assumption that the exponents associated with gender are floating and they need to be linked to a temporal unit in order to be phonetically interpreted. The structural position of the underspecified vowel attracts those features, as shown below for the feminine noun *casa* 'house':

(3)     *casa* 'house'[3]

```
                         n
                 _____|_____
                /                 \
            √cas                    n
                            _____|_____
                           /                 \
          n_[+fem]↔[+low]             Theme
                       ⸌                   |
                        ⸌                  V
                         ⸌_____⸍
```

---
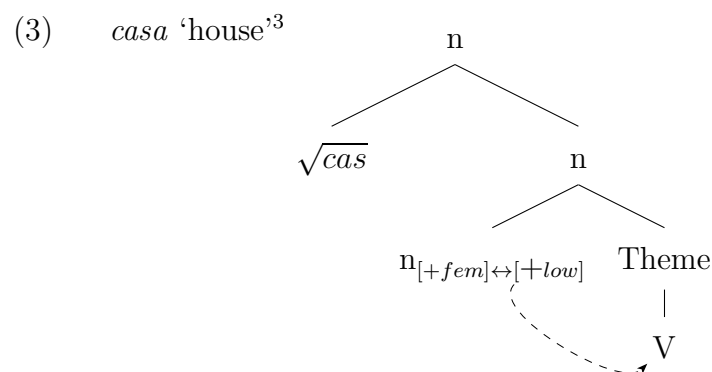
[2]In what follows, I analyze all the different vowel endings of Catalan nouns and adjectives (see tables 7.1 and 7.2 above) as reported in morphological studies of Standard Catalan, based mostly on Central varieties (e.g., Mascaró 1986; Clua 2002; Lloret 2013). Not all the forms in 7.1.1 and 7.1.2 were reported during the interviews, but, by extension, the inflecional endings of Standard Catalan are nevertheless assumed for Pallarese (there is not a lot of dialectal variation regarding the shape of these morphological markers, besides, obviously, the different patterns of vowel reduction).

[3]For simplification purposes, in this chapter I do not use phonetic symbols in spelled-out roots but orthographic conventions instead.

In chapter 5 it was argued that, in 3rd person singular *l*-clitics, the feminine always surfaces as *-a* due to a high-ranked constraint PARSE[+LOW] that dominates DEPLINK. Recall that these two constraints are in conflict with each other because PARSE[+LOW] demands the parsing of [+low] and, consequently, its phonetic interpretation, whereas DEPLINK, which penalizes the insertion of new association lines, fights against the parsing of phonological features. The constraint ranking is thus crucial for generating the correct default pattern in the feminine (in (5)) and also in the masculine (in (4)), where PARSE[+LABIAL] is dominated by DEPLINK and the phonological features that correspond to /o/ cannot be pronounced. That is, the fact that DEPLINK outranks PARSE[+LABIAL] favors candidates that do not epenthesize association lines, which implies that the features associated with the masculine remain floating and, after Stray Erasure, they are absent in the phonetic output.[4] Therefore, the final unmarked outcome for the masculine consists of a noun without overt phonological exponence:

(4)     *Masculine regular noun:* nen *'child'*

| [+labial] /n ɛ n V/ | $\sigma$STR | DEP-[−BK] | DEP-LK | PARSE[+LAB] |
|---|---|---|---|---|
| a. ☞ [+labial]<br>n ɛ́ n | | | | 1 |
| b. [+labial] \|<br>n ɛ́ n o | | | 1W | L |
| c. [+labial]  [-back] \|<br>n ɛ́ n e | | 1W | 1 | 1 |

---

[4]As stated in previous chapters, [+labial] is a shortcut for all the features that correspond to the vowel /o/.

(5)    *Feminine regular noun:* nena *'child'*

| [+low]<br><br>/n  ɛ  n  V/ | σStruc | Dep-[−bk] | Parse[+low] | Dep-Link |
|---|---|---|---|---|
| [+low]<br>│<br>a. ☞  n  ɛ́  n  a | | | | 1 |
| [+low]<br><br>b.    n  ɛ́  n | | | 1W | L |
| [+low]    [-back]<br>│<br>c.    n  ɛ́  n  e | | 1W | | 1 |

Masculine and feminine nouns, as shown above, differ in their exponence due to the unequal ordering of their parsing constraints with respect to DepLink, which favors the linking of [+low] in the feminine but not of [+labial] in the masculine (Parse[+low]>>DepLink>>Parse[+labial]), thus generating the expected outputs *nen* and *nena*. In both tableaux there is a candidate with the features of the vowel *-e* that occupy the V-slot of the theme position, but it will never be the winner as it has to epenthesize those features (in the event of phonological conflict, the features used to improve syllable structure are those associated with gender, as will be shown in the next section). This is only a solution in cases like the plural clitic *ls/les* analyzed in section 5.1.2, where phonotactics force the insertion of default features in the theme position because there are no gender exponents available. In addition, I obviated a candidate with full epenthesis because it would be ruled out by Dep-V, which penalizes not only the insertion of features but also the creation of new skeletal positions.

In chapter 5 it was mentioned that a set of masculine nouns ending in a sibilant in the singular insert the masculine exponent *-o* in the plural to avoid the contact of two sibilants. These regular nouns behave as expected in the singular, i.e., without any overt exponent, as shown in (6) (following the same pattern of (4)):

(6)     *Regular exponence:* gos *'dog'*

|  [+labial]<br><br>/g  o  s  V/ | $\sigma$Str | Dep-[−bk] | Dep-Lk | Parse[+lab] |
|---|---|---|---|---|
| [+labial]<br><br>a. ☞  g  ó  s |  |  |  | 1 |
| [+labial]<br>\|<br>b.    g  ó  s  o |  |  | 1W | L |
| [+labial]  [+back]<br>      \|<br>c.    g  ó  s  e |  | 1W | 1 | 1 |

The difference with regard to other regular nouns lies in the plural. While the word *nen* 'kid' simply adjoins the plural morph to the root, the word *gos* needs to insert the masculine exponent *-o* to avoid an OCP sequence \**goss*. The morphological structure of the plural in (7) sheds light on the mechanisms that motivate the use of *-o* as a rescue strategy:

(7)     *gossos* 'dogs'
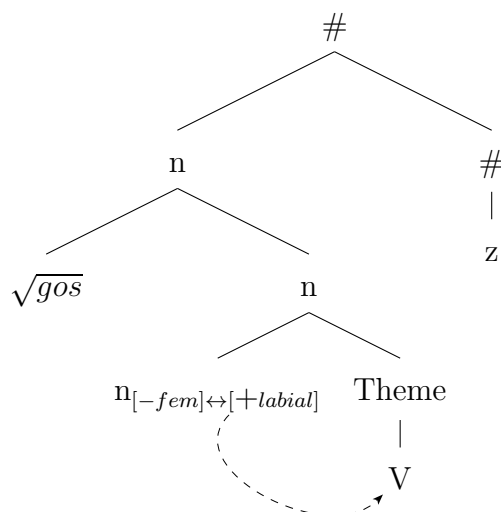


The featural exponents that correspond to [–fem] are realized as part of Vocabulary Insertion in regular nouns. However, in (6) it is shown how DepLink penalizes the insertion of association lines that connect the features of *-o* to the V-slot of

the theme position, which remain thus floating. In the case of the plural, the OCP contact between the root and the plural sibilants forces the use of a vowel, that is precisely *-o*, which remains floating in the singular but is linked in (8) due to phonotactic requirements:

(8)  *Morphological epenthesis in OCP-sibilant cases:* gossos *'dogs'*

| [+labial]<br><br>/g  o  s  V  +  z/ | σStr | Dep-[-bk] | Dep-Lk | Parse[+lab] |
|---|---|---|---|---|
| [+labial]<br>\|<br>a. ☞  g  ó  s  o  s | | | 1 | |
| [+labial]<br><br>b.    g  ó  s  s | 1W | | L | 1 |
| [+labial]    [-back]<br>\|<br>c.    g  ó  s  e  s | | 1W | 1 | 1 |

The high-ranked constraint σStruc is decisive in the final outcome of the plural.[5] It dominates DepLink and therefore candidate b. is ruled out although it satisfies the latter constraint. Candidate a. is the winner because it incurs the minimum possible violations. That is, the pronunciation of the plural form is only possible after the insertion of a vowel, so some association lines must be epenthesized anyway. As the feature [+labial] is already present in the input, it is preferable to link it to the V-slot than epenthesizing [−back] and linking it also to the V-slot, because, in any possible scenario, candidate c. has to epenthesize association lines.[6] The situation

---

[5]In (16), in chapter 5, it was argued that σStruc is a cover constraint cosisting of SonSeq and *ConsonantalNucleus. While the contact between two sibilants in (8) is better expressed by the constraint OCP-sibilant (see Bonet et al. 2007 for the treatment of OCP cases like *gossos*), I have used the σStruc constraint to maintain the same ranking for all cases in the nominal domain, as all the phonotactic problems have to do with syllable structure. SonSeq penalizes falling sonority onset clusters and rising sonority coda clusters, so it can be understood in a broad sense as not permitting the same sonority in coda clusters either (and thus ruling out candidates like *\*goss* with two sibilants).

[6]A candidate with pure epenthesis is not included by the same reasons outlined earlier in this section.

thus parallels the alternation *l/lo* in the masculine singular definite article and 3rd person accusative clitic, showing that morphology can work as a means to repair phonology. Therefore, the present analysis can cope with this set of plurals (which have been a puzzle in Catalan studies since the early days of generative grammar) once the traditional classification of masculine nouns with right-edge epenthesis has been discarded by the line of reasoning shown in 4.1.2.

Some other nouns can also be grouped within the 'morphological epenthesis' pattern. Wheeler (1979), Viaplana (1991) and Clua (2002) consider that words like *gerro* 'vase', are phonologically and *lexically* conditioned, because a sequence of two underlying taps —according to their analysis— is not permitted word finally but, nevertheless, the expected epentethic vowel [ə] (or [e] in Pallarese) is not found. The same holds for nouns ending in a non-tolerated coda cluster plus a final -*o* (i.e., –*CCo*, like *musclo* 'mussel', for example), which are also analyzed (besides the above mentioned studies) in Lloret & Viaplana (1992). My proposal can easily accomodate these nouns, which behave like *gossos* in (8). That is, as σSTRUC outranks DEPLINK, outputs like *musclo*, with [+labial] linked to the the V-slot of the theme position, will be preferred over candidates whose final coda cluster is non syllabifiable, i.e., *\*muscl*. An output like *\*muscle* is discarded like *\*gosses* for the reasons discussed above.

### 7.1.1.2  Lexicalized theme vowels

The autosegmental model proposed in this work for gender inflection can account for the regular behavior of Pallarese Catalan nouns, i.e., it can easily explain the default -[a] feminine nouns and -[Ø] masculine nouns. However, it has been shown that the picture is far more complicated than this, since there are many other endings that are used as inflectional markers. The basic idea behind the present analysis is that most of these endings are lexicalized and occupy directly the theme position in the lexical entry.[7] That is, these nouns bear fully specified theme vowels (which are not linked to syntactic gender), and thus the phonological exponents associated with either masculine or feminine cannot end up in the V-slot of the theme position — precisely because the placement of the underspecified vowel is occupied by a vowel with place features.

This idea is similar to Bermúdez-Otero (2013)'s analysis of the Spanish nominal system, as I am proposing that nouns not belonging to -*a* feminine and -*Ø* masculine regular cases fall out of the featural affixation analysis and are instead realized in

---

[7]Some other vowel endings that have been traditionally considered inflectional markers are assumed to be part of the root, as -*us* in *virus* 'virus'.

one single step with their specified theme vowel, i.e., they are realized as stems. This is illustrated in (9) with the feminine word *base* 'base':

(9)     $\sqrt{\text{BASE}}$-*n*-Th $\leftrightarrow$ /baze/

We can see in (9) that the whole nominal structure is realized by the exponent /baze/. Classic DM rejects entries like (9), which simultaneously realize multiple morphosyntactic terminals, but other versions of DM or other lexical-realizational frameworks like Nanosyntax (Starke, 2009) do allow them.[8] This is the proposal found in Caha (2009), Siddiqi (2009), Svenonius (2012) and Merchant (2015), who argue that phrasal spell-out/spanning is more adequate than single terminal insertion because it can get rid of morphological operations such as Fusion, which are problematic in DM (they need to be done at the postsyntactic morphological component and, therefore, there is still a distinction between syntax and morphology, something that DM aims to eradicate).[9] For the present purposes it is not relevant which theoretical framework is most adequate as a general theory of language; what really matters is that all of them allow the postsyntactic realization of complex structures, and therefore the Pallarese data analyzed here can be accomodated into any of these proposals.

It needs to be clarified that, even if the lexical item /baze/ spells out a whole nominal structure, the vowel -*e* realizes specifically the theme position. That is, *Th* matches /e/, and therefore the lexical entry in (9) must be understood as (10):

(10)     $\sqrt{\text{BASE}}_\alpha$-*n*-Th$_\beta$ $\leftrightarrow$ /baz$_\alpha$–e$_\beta$/

I follow Bermúdez-Otero (2012, 2013, 2016) in assuming that morphosyntactic information and phonological exponents are coindexed ("At a minimum, this will involve coindexing each phonological piece with the syntactic node it exposes; (...) I suggested above that such coindexation is present in the phonological input only" Bermúdez-Otero 2012, p. 52), which implies that the phonological realization of a noun reflects its constituent structure. Since this coindexation is present only in the input, input-output phonological constraints cannot refer to morphosyntactic constitutents, which are invisible for outputs.

An alternative option in a classic DM approach would be to assume that there

---

[8]See Caha (2016) for a discussion on the differences between DM and Nanosyntax regarding lexical insertion.

[9]Caha (2009, p. 51): "Further, I show that the architecture of grammar can be simplified if phrasal spell out is adopted. In particular, the need for certain morphological operations proposed within the framework of Distributed Morphology (Halle and Marantz 1993) disappears. Since these operations are assumed to operate in a separate module of the grammar, the module itself becomes emptier and we make important steps towards its elimination."

are specific Vocabulary Items for non-default theme vowels, which would have lists of nouns (or adjectives) specified in their context for insertion, as shown in (11-a) (cf. (11-b)):[10]

(11)    a.    Th ↔ e /{$\sqrt{\text{BASE}}$, $\sqrt{\text{CLASSE}}$, $\sqrt{\text{PARE}}$, ...}___
         b.    Th ↔ V

The Vocabulary Item in (11-a) shows that feminine nouns like *base* 'base' or *classe* 'class' and masculine nouns like *pare* 'father', which deviate from the regular pattern of feminine/masculine exponence, must be listed to get the correct inflectional ending. The Subset Principle (Halle, 1997) guarantees that roots specified in the Vocabulary Item of *-e* win the competition for insertion, with V being inserted in regular nouns instead, i.e., (11-b). While this solution works from a technical point of view, it "blurs the distinction between storage and computation by embedding lists within rules", in Bermúdez-Otero (2013, p. 31)'s words. That is, having lexical entries for theme vowels like those of (11-a) means that lexicalized structures are realized by rules. This implies, in fact, that there is no difference between regular and non-regular theme vowels (one Vocabulary Item is simply more specific than the other). Spelling out whole morphosyntactic trees reflects the exceptionality of these marked structures and retains the idea of a regular process of theme vowel realization, since there is only one specific Vocabulary Item for theme vowels, i.e., (11-b).

Although, as noted above, the realization of nouns with theme vowels in one single step is similar to Bermúdez-Otero (2013)'s proposal, he proposes stem-storage for *all* noun classes, even the default ones.[11] This is a crucial difference with respect to my analysis because it would imply that there is no general process of gender exponence for nouns like the ones presented in section 7.1.1.1. For this reason, the proposal defended in this thesis can be considered a hybrid between opposing theories of root and stem storage.

The advantage of storing theme vowels in lexical entries, as already discussed in section 1.1, is that we do not need to resort to diacritics on roots, which are problematic because they are 'alien' objects in the module where they appear, as Scheer (2012, p. 71) points out:

> A diacritic is a non-native object in module X: it is only used when
> information from outside of X is processed. It is absent from events that

---

[10]I would like to thank Eulàlia Bonet, Víctor Acedo-Matellán and Ricardo Bermúdez-Otero for discussion on multiple/single terminal insertion in DM.

[11]For Bermúdez-Otero, a *regular* Spanish masculine noun like *mano* 'hand' is stored as /man-o/, with the theme vowel already attached to the root.

do not appeal to extra-Xal information.

With lexical items like (10), the need for diacritics disappears. Any proposal that aims to provide a full analysis of Catalan nominal inflection needs to specify non-regular nouns somehow, but storing theme vowels allows us to get rid of foreign objects that refer to phonological exponence in the morphosyntactic component (i.e., 'a non-native object in module X'); therefore, this option should be favored.

Even if theme vowels are considered to be part of lexical entries in some cases, it is undeniable that not all non-default endings can be treated equally, so their most remarkable differences will be examined one by one.

Perhaps the most striking fact of the analysis is the assumption that /o/ is the default exponent for the masculine, as this is in fact a marked exponent in terms of frequency (most masculine nouns, by far, have a -Ø ending). How can we treat, then, a marked exponent as the unmarked option for the masculine? The idea is that the association lines that connect the phonological features of /o/ to the skeleton are already stored in the lexical entry, as opposed to the regular default pattern, where no place features are linked to the theme position. A comparison of both structures is offered below:[12]
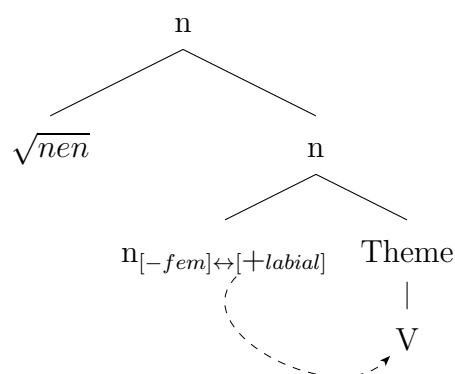
(12)    *nen* [nɛ́n] 'kid'              (13)    *nano* [náno] 'kid' [13]
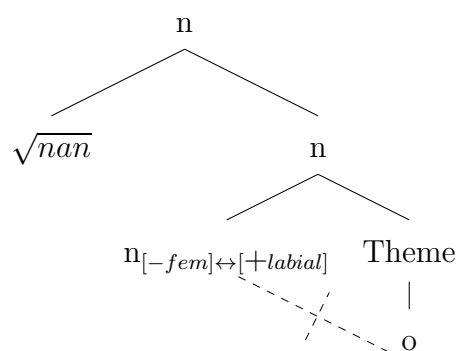


It is clear from the examples above that the structure for -Ø and -o masculine nouns is different. In (12), the regular process of assigning [+labial] as the default exponent for masculine nouns applies to the word *nen*. In the present analysis, these features do not end up in the V-slot of the theme position due to the aforementioned

---

[12]The few feminine words with an -o ending, like *moto* 'motorcycle' —which is, in fact, an abbreviation for *motocicleta*— or *soprano* 'soprano', also contain -o in the lexical entry, like *nano*. The gender specification in the morphosyntactic structure of these feminine nouns is obviously different.

[13]The broken straight line in (13) with a small line crossing indicates that the process of gender spell-out cannot apply in a regular way because there is a theme vowel specified in the lexical entry.

ranking DEPLINK >> PARSE[+LABIAL] in the masculine, which explains the asymmetry with regard to the feminine. In (13), by contrast, the vowel -*o* already occupies the theme position in the lexicon, thus preventing the generation of a regular word *\*nan* (in fact, this word exists in Catalan, but meaning 'dwarf' instead of 'kid'). One could argue that it is surprising that the default exponent for the masculine, -*o*, needs to be lexicalized. What is lexicalized, as mentioned, are the association lines that connect the featural specifications of -*o* to the timing of the morphological theme position. In both words *nen* —(14)— and *nano* —(15)— there is a place node with the features of the vowel -*o*, but in the regular case, *nen*, these features are simply spelled out by the morphological component without any connection to the phonological skeleton, while in *nano* the association lines are lexically stored:

(14)    *(Floating) featural affixation*        (15)      *Lexicalized association lines*


         [+labial]                                [+labial] [–cons]



              V                                         X

In (14) Vocabulary Insertion proceeds normally. First, the theme position is filled with an underspecified vowel (Th ↔ [–cons]), and then the phonological exponents for both masculine and femenine gender are spelled out, although they remain floating. For cases like (15), by contrast, the theme position already contains the place features of -*o* (recall that the whole nominal structure is phonologically spelled out in one single step with the theme position). It is important to note that, for masculine nouns like *nano* in (13), -*o* is still considered a gender marker because it shares exactly the same phonological features of the lexical entry for the masculine, i.e., [–fem]↔[+labial,–high].

In non-plural contexts, all proposals have to stipulate the presence of -*o* either in the lexical entry as part of the stem (as I do) or as a diacritic in the root, but my analysis has the advantage of generating -Ø masculine nouns and the morphological epenthesis cases in plurals analyzed in section 7.1.1.1 in a uniform way. If we treat -Ø and -*o* as different allomorphs, the interdependence of the two phenomena are more difficultly formalized, as traditionally, except for Lloret & Viaplana (1992) and Bonet et al. (2007), the appearance of the masculine morph -*o* in phonologically conditioned environments has been vaguely analyzed. Most studies say it is a morphologically conditioned epenthesis but do not explain *why* this is the case and which mechanisms rule out regular phonological epenthesis. With floating features in the input we can

account for the productive morphological epenthesis cases that we see in OCP-avoidance plurals like *gossos* 'dogs' (sg: *gos*; pl: *\*goss*) or in the singular masculine clitics exposed in 5.1 without the need to resort to allomorphy in lexical entries, as in Bonet et al. (2007). Furthermore, besides the unitary treatment of the masculine nominal morphology, we can maintain strict modularity between morphology and phonology.

Masculine nouns ending in *-a* behave like *-o* nouns: the theme vowel is incorporated into the theme position in the lexical entry, as in (16):

(16)     *drama* 'drama'

$$n \;\rightarrow\; \left[\; \sqrt{dram} \quad n \;\rightarrow\; \left[\; n_{[-fem]\leftrightarrow[+labial]} \quad \text{Theme} \;\mright| \; \text{a} \;\right]\right]$$
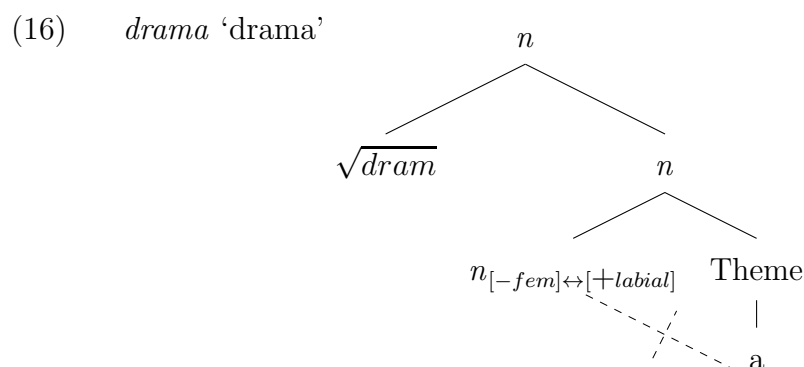
The main difference is that we do not find the masculine morph *-o* in the theme position but the unmarked inflectional ending for [+fem] instead, although this does not pose any problem because, as already argued, gender is an inherent property of nouns, not of theme vowels.[14] Therefore, the fact that *-a* is contained in the lexical entry blocks the regular process of default masculine exponence, which would generate the word *\*dram*. It is important to understand that the fact that vowels are stored in the lexical entry of these words does not imply that the regular process of inflection ceases to exist. In fact, in (16) it is shown that this process tries to apply but, as indicated by the broken straight line, the features associated with [–fem] will never end up in the theme position for two reasons:

1. The constraint PARSE[+LABIAL] is dominated by DEPLINK. This means that [+labial] —when it is linked to the regular process of inflection— will never surface (only when there is a phonotactic problem, as we have already seen). Moreover, the features associated with *-a* in *drama* are already contained in the input, which implies that DEPLINK cannot penalize the use of this vowel because its association lines are stored in the lexical entry (it only penalizes the insertion of association lines that are not present in the input). Thus, the candidate with the same theme vowel of the input will be preferred.

---

[14]The vowel *-a* in *drama* has nothing to do with gender; it is just a theme vowel that happens to share the phonological features associated with [+fem].

2.  A candidate with [+labial] features attached to the vocalic slot of the theme position (i.e., *\*dramo*) can be generated by GEN. Besides the fact that, as previously mentioned, DEPLINK disfavors the insertion of association lines for the features associated with [–fem], IDENT(FEATURE) will favor candidates with the same input-output featural composition.[15]

A more complicated analysis regards invariable nouns such as *atleta* 'athlete', used both for masculine and feminine referents.[16] The feminine default exponent is *-a*, so in the feminine case, according to my proposal, it behaves as a regular noun:

(17)    *atleta* 'athlete' (fem.)
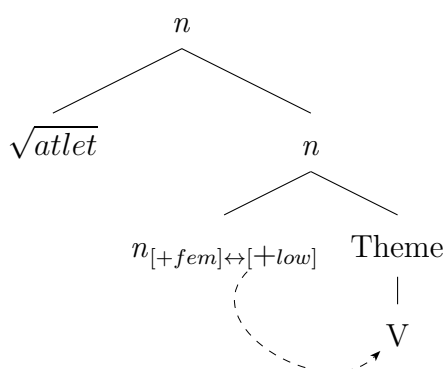


If all non-default endings are lexicalized, as I am assuming, the vowel *-a* should be stored in the theme position of the masculine noun instead:

(18)    *atleta* 'athlete' (masc.)
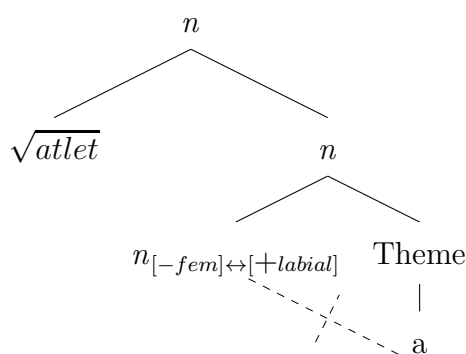


Therefore, there is an asymmetry in the storage of this group of nouns, as they behave regularly in one gender (the feminine in the case of *atleta*) but not in the other, where the theme vowel is already contained in the lexical entry.

---

[15]In masculine *-o* nouns like *nano*, where the lexicalized vowel corresponds to the masculine default exponent, there is no tension between the theme vowel of the lexical entry and the regular process of masculine inflection, as their featural composition is the same.

[16]Epicenes such as *víctima* 'victim', used also for both masculine and feminine referents but with only one gender (feminine in this case), are quite special in the nominal system. They have an inherently specified syntactic gender (which does not correspond necessarily to biological sex) by which their phonological features are determined. As opposed to nouns such as *atleta*, with agreement in two genders, they show agreement just in one gender, and thus we find sentences like *una petita víctima* 'a small victim', even if the antecedent is masculine.

There is a set of nouns, that I label *-e* stems importing the terminology from Spanish noun inflection studies, that bear *-e* as a theme vowel (pronounced as -[ə] in Central Catalan due to vowel reduction). In the masculine, these nouns are usually classified into two groups, one where *-e* must be a lexical specification, as in (19-a), and one where this vowel appears due to phonotactic requirements (i.e., it is considered epenthetic), as in (19-b):

(19)    a.    *pare* 'father'
        b.    *centre* 'center' (*centr)
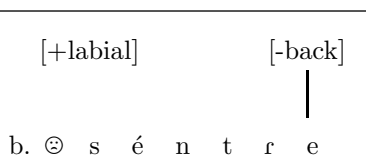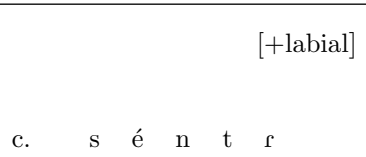
It is commonly assumed that the final vowel in (19-b) lies outside the morphological domain because it is epenthetic, i.e., ([centr]e), whereas in (19-a) the vowel has a morphological affiliation (it is inflectional), i.e, ([par-e]). This distinction has been assumed almost pervasively in the literature (see Mascaró 1986; Bonet & Lloret 1998; Lloret & Viaplana 1992; Wheeler 2005; Bonet et al. 2007 among others), but it is argued in this dissertation that these two groups should be conflated into one (see section 4.1.2 for an overview of epenthesis in Catalan, especially at the right edge of the word). The vowel in (19-b) was epenthetic in previous stages of the language, but I argue that right-edge epenthesis in Catalan nouns is not a synchronic process. If this idea is on the right track, the vowel must be in all cases lexicalized (both in *centre*, (20), and *pare*, (21)), as shown below:

(20)

```
              n
          ____|____
         /         \
    √centr          n
                 ___|___
                /       \
   n[−fem]↔[+labial]   Theme
              ⟍  ⟍      |
                ⟍  ⟍    e
```

(21)

```
              n
          ____|____
         /         \
    √par            n
                 ___|___
                /       \
   n[−fem]↔[+labial]   Theme
              ⟍  ⟍      |
                ⟍  ⟍    e
```
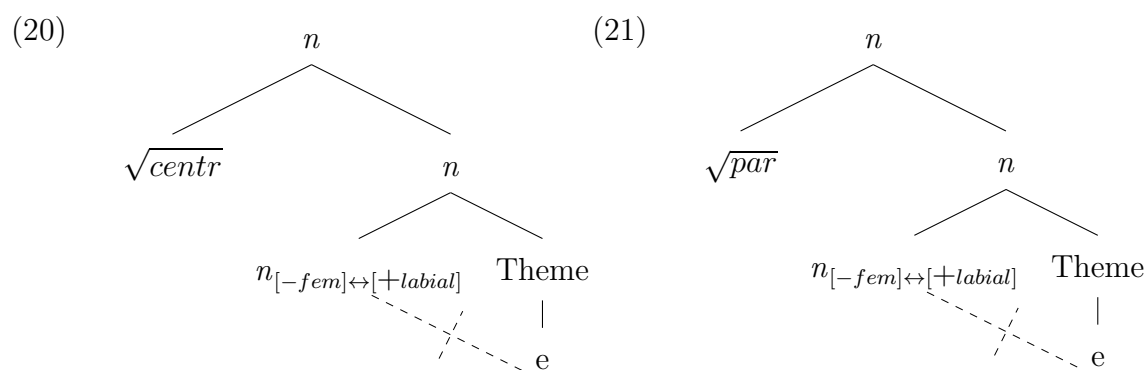
There is no distinction between (20) and (21), as both nouns belong to class *-e.* To my understanding, this is a simplification of the inflectional noun system of Catalan that groups what is usually divided into subclasses into one big class. This new classification is a consequence of the proposal, because an epenthetic analysis of *centre*-like words would overgenerate nouns with an *-o* ending. That is, according to the autosegmental inflectional analysis of regular masculine nouns, the exponents of [–fem], i.e., those associated with /o/, remain floating in the input and are in fact non-parsed because of the ranking DEPLINK>>PARSE[+LABIAL]. Only in case of

phonological conflict are they used to improve syllable structure, as in the 3rd person accusative clitic and definite article *l/lo*. My proposal predicts that the illicit coda cluster caused by historical apocope should be repaired (*contra* previous analyses) by linking the floating features associated with [–fem] (i.e., *-o*) to the theme position, as shown in (22):

(22)     *Morphological epenthesis in the masculine,*
         *centro (incorrect prediction)

| [+labial]<br><br>/s  e  n  t  ɾ  V/ | σStr | Dep-[−bk] | Dep-Lk | Parse[+lab] |
|---|---|---|---|---|
| [+labial]<br>│<br>a. ☞  s  é  n  t  ɾ  o | | | 1 | |
| [+labial]     [-back]<br>│<br>b. ☹  s  é  n  t  ɾ  e | | 1W | 1 | 1 |
| [+labial]<br><br>c.  s  é  n  t  ɾ | 1W | | L | 1 |

Nouns like *centre* are usually interpreted as belonging to the unmarked class -Ø, which implies that, in their morphological structure, they consist of just a root, being the right-edge vowel *-e* epenthetic. This classification, however, cannot be maintained with the autosegmental analysis of regular default nouns, as shown in (22). Candidate c. is ruled out because it violates σStr and does not respect the Minimum Sonority Distance established for Catalan. Candidate b., by contrast, allows the syllabification of the coda cluster, but it does so by epenthesizing [–back] at the expense of violating Dep-[−back], which dominates DepLink and Parse[+labial]. The place features of *-e*, though, end up occupying the theme position and thus the vowel becomes thematic, but it is anyway more parsimonious to use the [+labial] feature that corresponds to [–fem] to solve the phonological problem.[17]  Therefore, given this scenario, σStr would force the linking of the

---

[17]Again in this case, the morphological structure of the noun system predicts that a pure

vocalic features corresponding to masculine that are otherwise not associated due to the highly-ranked DEPLINK —cf. (4).[18]

For the 'proper' functioning of the Catalan noun system we must assume, then, that this apparently epenthetic vowel is already specified in the word's lexical entry. To do so could seem a stipulation, but it is not when we observe the epenthetic patterns of Catalan. If right-edge epenthesis ceased to work as a phonological process and it is not active in loanword adaptation, we have a strong argument to assume that the vowel *-e* is lexical in nature. Furthermore, any model has to deal with *-e* nouns like *pare* whose vowel must be specified in the lexicon, and the move here is just to add more masculine nouns to this category once we have gotten rid of the 'false' epenthetic class.

What we can save with the autosegmental analysis, though, is the very common OCP-avoidance masculine plurals that are rescued through the insertion of *-o*, the case of words like *gossos* 'dogs' , as well as nouns with illicit coda clusters that bear an *-o* ending, like *musclo* 'mussel' (*\*muscl*), mentioned earlier in this section. The case of *musclo* is exactly the same as *gossos* because the vowel *-o* is also used as a repair straegy, but previous studies could not deal with the differences in right-edge epenthesis between *centre* and *musclo*-like words, and simply had to stipulate that certain words are lexically marked to select an inflectional ending instead of a regular epenthetic vowel when required by phonotactics. By having floating features, we can explain different phenomena with one single mechanism if we simply distinguish between a regular process of vowel insertion under phonological conditionings (in fact, masculine gender exponence, i.e., *-o*) and lexicalized theme vowels at the right edge of the word (i.e., *-e*).

As for *-e* nouns, the feminine situation does not parallel the masculine cases, as there is no noun which can be analyzed under the epenthetic analysis. That is, while historical apocope (and ulterior epenthesis) affected masculine nouns, feminine *-a* nouns remained unaltered and there was no absolute final coda cluster that posed syllabification problems.[19] What we find, though, are lexical *-e* vowels, as in *mar*[e]

---

epenthetic vowel cannot be inserted because it creates a morphophonoloigcal mismatch. For this reason, a candidate with full epenthesis —features + skeletal position—, not included in (22), would be ruled out by DEP-V.

[18]Actually, this is what happens in Ribagorçà, a Northwestern Catalan dialect spoken in Northwestern Catalonia and La Franja area —an Aragonese Catalan-speaking strip adjacent to the Western border of Catalonia— where the group of nouns traditionally analyzed as epenthetic surface with an *-o* ending instead of *-e* (*negro* 'black', *centro* 'center' vs *negre*, *centre* in other varieties of Catalan; cf Sistac 1993).

[19]This is what would have happened in masculine/feminine pairs that have *-tr* coda clusters, for example, like *pediatr*[e]/*pediatr*[a] 'pediatrician' in Northwestern Catalan (*pediatr*[ə] in Central Catalan for both genders). In the masculine, the vowel from the latinate base disappears, thus yielding an ill-formed sequence *\*pediatr* that subsequently undergoes epenthesis. In the feminine,

'mother', *bas*[e] 'base' or *piràmid*[e] 'pyramid'. These endings must be, in any morphological model, specified in the lexical entry, as no phonotactic reason motivates their use and they deviate from the regular pattern that realizes feminine as -[a]. We must conclude, then, that their lexical representation, depicted in (23), is the same as in the masculine:

(23)     *base* 'base'
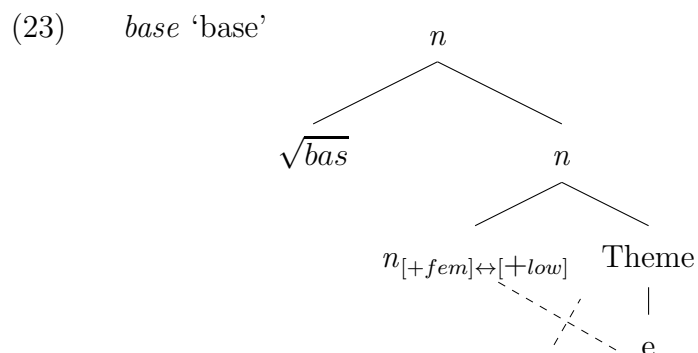


If the theme vowel is already contained in the lexical entry, the regular process of inflection is blocked and therefore the constraints that guarantee the final exponence in -*a* feminine nouns —i.e., DEPLINK and PARSE[+LOW], in short— are not active, the same that happens with lexicalized nouns in the masculine. This group, then, does not pose any major problems and it patterns like the other non-regular nouns analyzed until now.[20]

Something which is difficult to account for in my model are feminine nouns with a -Ø ending, such as *sal* 'salt', which should surface as *\*sala* if they followed the regular inflectional pattern of the feminine. Until now, all the masculine regular nouns have been analyzed as having a -Ø ending, a somehow confusing terminology, as actually these nouns have an exponent which is /o/ but it is not linked to the skeleton in the phonological component and, thus, it is not interpreted phonetically. It is more accurate to treat these set of nouns as consisting of non-pronounced exponents, but I group them as having a -Ø ending for exposition purposes. However, the feminine nouns which do not have an overt inflectional marker cannot be treated equally, as the problem is not phonological but morphological. That is, while in regular masculine nouns the non-linking of the floating features associated with [–fem] is phonological in nature (the Vocabulary Insertion process proceeds normally in the morphological component, associating

---

instead, the presence of the inflectional ending -*a* does not provide the context for epenthesis.

[20]The lexicalization hypothesis in this set of feminine nouns holds true for Pallarese and other Western dialects that do not reduce unstressed *e/a* to schwa. For speakers of eastern dialects like Central Catalan, for example, that produce outputs like *piràmid*[ə] or *mar*[ə], these nouns are as regular as *cas*[ə], and thus the vowel follows the regular pattern of inflection and does not need to be lexicalized.

certain syntactic values to certain phonological features), the non-presence of -*a* in the feminine cannot be explained under the same logic. The constraint ranking PARSE[+LOW]>>DEPLINK>>PARSE[+LABIAL] guarantees that the feminine exponent will be interpreted by the phonetics module, and thus a non-pronunciation scenario is discarded. The problem, then, is to determine whether this group of feminine nouns have a 'real' -Ø ending or not. As exposed in Trommer (2012), this is not a trivial question, as there are many different definitions of -Ø exponence that lead to a certain degree of confusion in morphological studies.

According to my analysis, words like *sal*, which deviate form the regular inflectional process, need to be stored in the lexicon with their corresponding theme vowel (or non-vowel in this case), as depicted in (24):

(24)     *sal* 'salt'
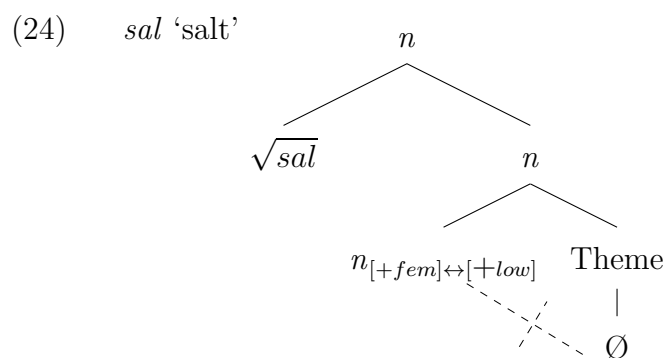


The problem is: what does -Ø really mean? Is it a morphological marker *per se* or just the absence of phonological features? Both options generate some problems that need to be analyzed in detail.

According to Trommer (2012, p. 328) one of the instances of zero exponence is what he labels 'Wu Wei Ø exponence'. This kind of exponence is actually the 'no exponence', as speakers do not create a Vocabulary Item for a specific bundle of features (an idea which actually goes against one of the implicit tenets of Distributed Morphology, named by Trommer the 'Axiom of Obligatory Exponence' —in the output of morphology, every head must be filled by at least one VI— and considered unnecessary). This idea is appealing for the feminine cases mentioned above, but it does not work because we need a Vocabulary Item corresponding to the theme position to account for all the regular cases (i.e., Th↔V) and also because structures like (24) are assumed to be stored in the lexicon as a whole (it cannot be viewed as a regular realizational process of the language).

Something that could work for the present purposes is Lampitelli (2014)'s analysis of theme vowels, which he considers lexical properties of the roots. This assumption implies that words that do not bear on overt inflectional ending do not project

a theme postition, as shown in (25) and (26) for Spanish:[21]
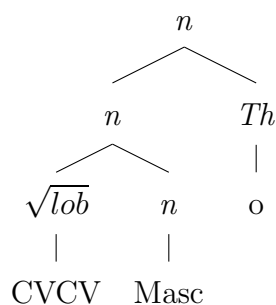
(25)    *lobo* 'wolf'                    (26)    *ciudad* 'city'
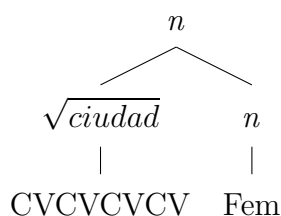
In (25), we see that the word *lobo*, which bears the default inflectional ending of the masculine in Spanish, -*o*, projects a theme position, while *ciudad*, which departs from the regular -*a* inflection for the feminine (*\*ciudada*), does not. Lampitelli's idea is simple: if there is no theme vowel there is no need for a theme position. His analysis is very adequate for the Italian nominal system because it can capture the asymmetries between the native lexicon and loanwords in plurals such as *film* 'film', which remain invariable even if a sequence *\*filmi* 'films' is totally acceptable (and actually preferable). The fact that the word *film* does not select a theme position prevents the realization of the plural.[22] While Lampitelli's proposal is attractive and could work for *sal*-like words, it must be rejected by the behavior of the already analyzed plurals like *gossos* 'dogs', where we clearly see that the theme position imposes a morphological solution to a phonological problem; if there were no theme position at all, we would expect a masculine plural *\*gosses*, with regular epenthesis, which is in fact what he proposes for Spanish in *ciudad-ciudades* 'city-cities' —the default epenthetic vowel of Spanish is -*e* as well, as in Pallarese.[23]

The problem we have now is twofold. On the one hand, we cannot treat the -Ø marker of these words as non-parsing of phonological features as we did in the masculine, because the feminine parses /a/ in the default cases. One option would be to assume that, in the lexical entries of these feminine words there is simply nothing in the theme position. In a sense, this could be understood as 'Wu Wei

---

[21]The morphological structure proposed by Lampitelli, which also includes a theme position (in certain cases), is similar to Oltra's, although the overall analysis of the nominal system is very different. The phonological model advocated there is CV phonology (Lowenstamm, 1996), but this is irrelevant in the present discussion.

[22]In his CV analysis, Lampitelli argues that Theme provides the CV unit for the plural in Italian. If there is no Theme projection, it cannot be realized.

[23]See section 7.1.2 for an analysis of the differences between masculine and feminine plural nouns and adjectives that end in a sibilant in the singular (e.g., *feli*[s] (masc./fem.) 'happy') will be treated

Ø exponence' following Trommer's terminology (although it would not be a rule of Vocabulary Insertion, it would simply apply in specific cases). The -Ø in (24), then, would be understood as silence, although the theme position would be maintained. The problem with this alternative is that silence or, in other words, the absence of phonological features, leaves the theme position open for the application of the Vocabulary Insertion rules that control exponence. That is, if there is nothing in the theme position, a V-slot will be inserted in it (Th↔V) and, after this, the exponents assigned to [+fem] ([+fem]↔[+low]). In this scenario, an output *sala* is wrongly predicted.

On the other hand, it is assumed in traditional studies that regular masculine nouns have a (real) -Ø ending (i.e., a zero morph), which is widespread in the morphology. Therefore, extending its appearance in the feminine in words like *sal*, even by the use of diacritics in the root, does not pose big problems. However, I am assuming that zero exponence in the masculine does not have to do with morphology but with phonology (i.e., with parsing), and thus the only instances of -Ø endings would correspond to this small set of feminine nouns. While this is not a truly satisfactory explanation (ideally, it would be better to get rid of all instances of zero exponence), there is no other choice but to assume that nouns like *sal* bear a zero marker in the theme position. If -Ø is a morphological entity, it can —and it does— block the application of the regular realizational process of the Catalan nominal system.[24]

An even more obscure situation is found in what has been labelled athematic nouns (in Spanish) in Bermúdez-Otero (2006a,b, 2007, and subsequent work). This terminology is transparent in Spanish because the vast majority of nouns bear an overt inflectional marker, but in Catalan the situation is different as the masculine is usually realized as -Ø. However, I will refer to this group of nouns as athematic because it captures, in a way, the idea of 'deviance', and I will extend it to other classes that do not bear the most common inflectional endings of Catalan seen until now, repeated here for clarity purposes:

- -*a* (default feminine, few masculine forms)

- -Ø (default masculine —which is in fact the non-parsing of /o/— and some feminine words)

- -*o* (lexicalized in some masculine nouns, sporadically in feminine nouns)

---

[24]The idea of zero morphs will receive further attention in section 7.1.2 when analyzing adjectives like *capaç* 'capable', which are invariable in gender in the singular but not in the plural (*capaços, capaces*).

- *-e* (lexicalized for both masculine and feminine nouns)

Therefore, any noun which patterns with a different ending from the ones mentioned above will be considered athematic, even if some of them apparently display theme vowels. Within this group, some distinctions need to be made as well, analyzed in detail below.

**7.1.1.2.1 Root allomorphy** There is a set of nouns, most of them learned words from Latin and Greek, that have *-Vs* endings (which were, originally, case markers), such as *virus* 'virus', *brindis* 'toast', or *cosmos* 'cosmos' . They have a special behavior because they maintain their ending in some derivatives, like diminutives, but not in others, e.g., *viruset* 'small virus' but *víric* 'viral'.[25] I follow Bermúdez-Otero (2007) in considering that these cases have to be understood as root allomorphs (i.e., *virus-, vir-*), because otherwise it is difficult to explain their behavior. Therefore, we have to assume the following structures:

(27)    *virus*
        'virus'
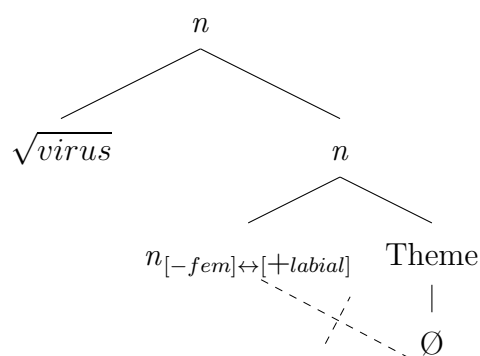
(28)    *víric*[26]
        'viral'
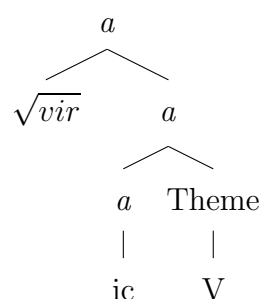
[26]*Víric* is the masculine form of the adjective (feminine = *vírica*), and it inherits its gender features by agreement with a noun; *cultiu víric* 'viral culture' (masc.) vs. *malaltia vírica* 'viral disease' (fem.).

(29)    *viruset* 'small virus'
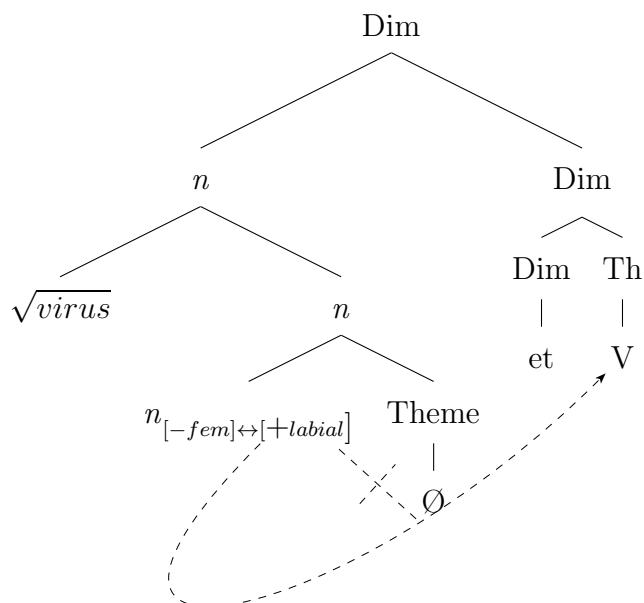


In (27), it is shown that *virus*, a masculine noun, is stored as having a zero morph (i.e., it behaves like the feminine noun *sal*). If *virus* is the whole root, one could argue that these nouns should be treated as regular -Ø masculine nouns. That is, they enter the derivation as any regular noun (e.g., *nen* 'kid', shown in (12)) and they do not end up as *\*viruso* due to the constraint ranking that favors the non-parsing of -*o*. The problem in assuming this lies in their plural formation. The root *virus* ends in a sibilant, which implies that it should be pluralized as *gossos* 'dogs' by linking -*o* to the skeleton in order to avoid the OCP effect. We find instead a plural *virus* 'viruses', with adjunction of the plural morph /z/ and degemination. This shows that it cannot be treated as a regular masculine noun and, in fact, the degemination analysis seems to be the most adequate as it captures the asymmetry between *gos* and *virus*-like words regarding plural formation. If we consider *virus* to be stored with a -Ø morph in the lexical entry, it is clear that the plural cannot be *\*virusos* as the features of [–fem] cannot end up linked to the theme position because it is blocked precisely by the -Ø morph. The adjunction of the plural morph /z/ to the sibilant of the root leaves no other option than degeminating and, as a consequence, the singular and the plural have the same exact shape.

If we adopt Bonet et al. (2007) analysis for this set of nouns (recall that, for them, all allomorphs are contained in the lexical entry, i.e., /virus{Ø>u>ə}/), the problem that we face is that words like *gos* and *virus* cannot be distinguished.[27] The winning candidates bear both a -Ø ending due to the action of the constraint PRIORITY, which favors the -Ø allomorph. When the plural morph is adjoined to

---

[27]The dialect analyzed in Bonet et al. (2007) is Central Catalan, hence the different inflectional endings.

the base noun, the constraint ranking correctly generates a plural like *gossos*, but it also predicts the plural *\*virusos* (wrongly ), as shown in (30):[28]

(30)      *\*virusos* (inorrect prediction)

| /biɾus-{Ø>u>ə}-z/ | OCP | Dep-V | Priority |
|---|---|---|---|
| a. ☞ bíɾusus | | | 1 |
| b.    bíɾusəs | | | 2W |
| c.    bíɾuss | 1W | | L |
| d.    bíɾus̲ə̲s | | 1W | L |

The overgeneration of the plural with the inflectional allomorph *-u* (candidate a.) is due to the assumption that both *gos* and *virus* have the same morphological structure, which consists of a root plus the stored allomorphs in the lexical entry. Instead, under my analysis, *virus* is lexicalized as having a *-Ø* ending in the theme position. With this modification we can easily capture the difference between the plurals *gossos* and *virus*, as the blocking effect of the *-Ø* ending forces the degemination solution after the contact of the plural morph (epenthesis is a last resort strategy at the right edge of the word, and as such, it is avoided in this kind of plurals, i.e., *\*viruses*).[29] It is important to emphasize that the *-Ø* morph is always the default for the masculine in Bonet et al. (2007) and, as such, it cannot be specified in the root in the form of a diacritic (only roots are stored in the lexicon in their model), as in nouns that have an *-o* or *-e* ending, such as *mosso* 'guy' (i.e., moss$_o$) or *pare* 'father' (i.e., par$_e$). They cannot deal with the different behavior of *gos* and *virus* because they do not contemplate the idea of lexicalized structures, something that is allowed in my analyisis. Moreover, they consider right-edge epenthesis to be a regular process of Catalan and, given this assumption, a plural *\*viruses* would always be preferred over *virus* because there is no fusion between the root and the plural sibilants.

---

[28]This is a simplified tableau for OCP-sibilant plurals. See Bonet et al. (2007, pp. 920-924) for more details.

[29]However, we will see later on in 7.1.2 how we actually find an epenthetic solution in the plural of feminine adjectives such as *felices* 'happy' (with a final sibilant in the singular, *feliç*).

Regarding the theme position, the examples in (28) and (29) show that it is filled with an underspecified vowel in both the diminutive and adjective. Once the suffixes in *víric* and *viruset* have been attached to their bases, the word becomes regular. There is no overt inflectional marker in either case, but we need to carefully differentiate between the absence of the theme vowel in *virus* and *viruset/víric*. In the former case, it is a specification of the lexical entry, whereas in the latter it is due to the non-linking of the vocalic features of the masculine default vowel *-o*. The exceptionality of the word *virus* fades away when confronted with a regular derivational process through the adjunction of the diminutive *-et*, that behaves regularly in the inflectional process.

It is important to keep in mind that derivational suffixes (such as *-et* or *-ic* here, but almost all others) behave as expected according to the inflectional regular process advocated in this work, which implies that most of the time their realization for gender is *-Ø* in the masculine and *-a* in the feminine. In (28) and (29) the gender of the base (in the case of the adjective *víric*, the noun it modifies) specifies the use of the masculine suffix (there are masculine-feminine pairs *-et/-eta* and *-ic/-ica*), and the theme position of the suffixal head just inherits the phonological features associated with [–fem], as shown in the morphological structure with the underspecified V.

In some other cases the suffix carries its own gender specification and modifies the word according to it. Take as an example the masculine word *portal* 'hallway', which derives from the feminine noun *porta* 'door' after the attachment of the sufix *-al*:[30]

(31)  El        portal      és gran
      DEF.MASC.S hallway.MASC.S is big.MASC.S
      'The hallway is big'

As shown in (31), gender agreement is masculine and there is no overt inflectional marker in the noun (as expected for the masculine), i.e., *portal*[Ø]. This clearly shows that the sufix *-al* must bear [–fem] features that modifiy the root it is attached to because otherwise the word would take the feminine features of the base *porta* 'door'.[31] Therefore, we must posit the following morphological structure for *portal* in (32):

---

[30]The suffix *-al* always creates masculine nouns. Other derivatives from feminine bases are, among others, *casal* 'dynasty', *didal* 'thimble', *finestral* 'large window' or *camal* 'trouser'.

[31]This is an argument against gender features in roots. See section 4.2.2.

(32)   *portal* 'hallway'
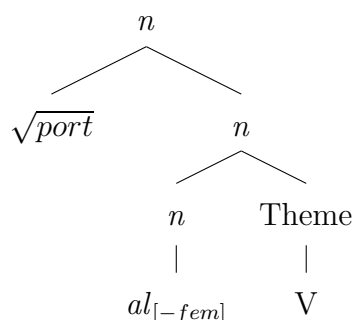


Regardless of where gender is stored, the phonological spell-out process will be the same. In most cases, as in *portal*, we will find the unmarked inflectional pattern in derived nouns, which, as we said, implies -Ø for the masculine and -*a* for the feminine. For some suffixes, though, we have to assume that their structures are lexicalized, as it happened with the vowels of non-derived nouns. This is the case for the suffix -*or*, which produces feminine words like *escalfor* 'heat' (a deverbal from *escalfar* 'to heat'). As a feminine noun, we would expect the output *\*escalfora* if the inflectional autosegmental process were regular. As this is not the case, we must assume that it is stored as $\{or\text{-}\emptyset\}_{[+fem]}$, with a -Ø inflectional ending in the theme position, the same assumption that has been made for feminine words that deviate from the unmarked pattern like *sal* 'salt'.

**7.1.1.2.2   Pseudoplurals (and pseudofeminines?)**   Some other nouns (apart from the ones mentioned previously), which are very rare and, most of the time, of very low frequency, have an -*s* ending as well (or -*es*). Lloret (2015) gives the following words as examples of this class:

|      |            | **Common nouns** |            | **Proper Names** |                      |
|------|------------|------------------|------------|------------------|----------------------|
| **-es** | *àlies*     | 'alias'          | *Carles*    | 'Charles'        |                      |
|      | *atles*     | 'atlas'          | *Elies*     | 'Elijah'         |                      |
| (33) | *mecenes*   | 'patron'         | *Hèrcules*  | 'Hercules'       |                      |
|      | *pàncrees*  | 'pancreas'       | *Sòcrates*  | 'Socrates'       |                      |
|      | *galimaties*| 'gibberish'      |            |                  |                      |
| **-s**  | *temps*     | 'time'           | *Àngels*    | 'Angels' (literally) |                  |
|      | *plebs*     | 'plebs'          | *Dolors*    | 'Pains' (literally)  |                  |

Some of these nouns are more common than others, but the crucial point is that not all of them have the same behavior; in some cases they retain the -*s* in the derivatives whereas in others the regular pattern of inflection is found. These nouns are very marked in terms of regularity and speakers are always very hesitant about

the 'correctness' of the output forms they produce in derivation, but nevertheless the differences they show shed light on their morphological structure.

Within the *-es* group, a word like *Carles* is analyzed as follows, with a lexicalized theme vowel *-e* and a pseudoplural morph /z/ (also lexicalized):[32]

(34)    *Carles* 'Charles'
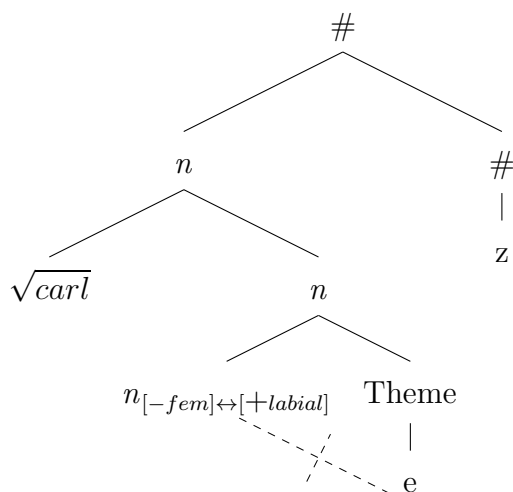


The *-s* of the noun is reinterpreted as a plural marker, although the noun counts as singular for agreement purposes (*En Carles és simpàtic* 'Charles is nice'). However, when speakers make diminutives from this base they can analyze it in two different ways: either they retain the *-s*, using *-et* as an infix, as in (35) (*Carletes* 'little Charles'), or they apply the regular suffixation procedure to the stem —(36) (*Carlet* 'little Charles')— obviating the pseudoplural morph. The suffix -{*et/-a*} is spelled out like regular nouns in Catalan, i.e., with no overt exponent in the masculine and *-a* in the feminine (the floating features associated with gender are attached to the V-slot of the theme position according to the general phonological ranking). Given the morphological structure [[carl-e]–et-V], the diminutive surfaces as in (36):[33]

(35)    carl<**et**>–e–s

(36)    carl–et

Lloret (2015) even provides another form *Carlets*, which is totally unacceptable in my idiolect. For the set of speakers that accept this form, it should be analyzed as (36) but including a pseudoplural morph, as in (37), which is a particularly odd

---

[32]This analysis is the same as the one given by Bermúdez-Otero (2006b) for the Spanish counterpart *Carlos*, being the only difference the quality of the theme vowel.

[33]In (36) the theme vowel *-e* is already contained in the lexical entry. In diminutive formation, this vowel is deleted by the stem-final vowel deletion rule posited in chapter 4 following Bermúdez-Otero (2007) (and subsequent work): [[carl-**e**]-et] > carl-et. This holds for all nouns and adjectives that contain a lexicalized theme vowel, which disappears in derivational contexts.

output (if the noun is already 'regularized' with the diminutive suffix, it should not retain the plural marker in the morphological structure: [[[carl-e]–et-V]–s]):

(37)    carl–et–s

As Bermúdez-Otero (2007, p. 247) points out "suffixation results in the loss of the pseudoplural ending: as an inflection, the /-s/ cannot occur on the left of a stem-based suffix, and the suffix in turn blocks its appearance on the right." He assumes that the Spanish diminutive *Carlitos* (from Carlos 'Charles') patterns like (35), with an infix, and forms like *Carlote* 'Carlos.AUGM' make use of the suffix, instead, and thus the non-presence of the pseudoplural ending. *Carlets* should be possible only as a real plural, not as a pseudoplural (what he assumes for evaluatives like *Carlotes*), and therefore the form in (37) is definitely an unexpected output. Be that as it may, the patterns shown above for these productions —(35) and (36)/(37), i.e., infixation or suffixation— follow Bermúdez-Otero (2006b, 2007)'s analysis of Spanish nouns and their derivatives.

It should be noted that speakers consider either *Carletes*, *Carlets* or *Carlet* very strange forms. The most common diminutive for *Carles* is actually the Spanish form *Carlitos*, which is obviously lexicalized. The "impossibility" of these forms has to do with the pseudoplural ending. As pointed out by Bonet (p.c.), it is probably the contradiction between the morphological requirement for number to be the peripheralmost category and the syntax-semantics incongruency of a plural singular which really creates the problem, as the diminutive must be spelled out before number but, at the same time, the plural morph is already part of the base.

Some other forms different from *Carles* do not accept the infixation procedure and only allow the regular diminutive suffixation. This is what we find in pseudoplural forms such as Sòcrates *'Socrates'* or Hèrcules 'Hercules'. While these nouns bear as well a lexicalized theme vowel *-e* and are also considered pseudoplurals (as (38) shows), their only possible diminutives are *Socratet* 'little Socrates' —(39)— and *Herculet* 'little Hercules':

(38)    *Sòcrates* 'Socrates'
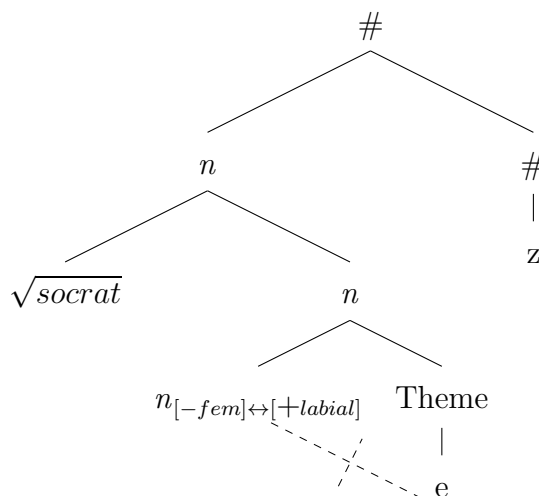


(39)    [[socrat-e]–et-V]

Therefore, *Socratet* behaves as *Carlet* in (36), which implies the addition of the diminutive suffix to the stem and the regularization of the word, that loses the plural marker. By contrast, another learned word from Greek and Latin, *mecenes* 'patron' (as well as *Elies* 'Elijah' or *Maties* 'Matthew'), maintains the *-es* ending, *mecenetes* (*\*mecenet*). *mecenetes* behaves like (35), which infixes *-et-* and thus retains the pseudoplural character, i.e., mecen<**et**>–e–s. Bermúdez-Otero (2006b, p.303) argues that this is due to a subcategorization requirement of the infix *-it-o/- a* to select for bases that contain the theme vowels /o/ or /a/. While this is true for Spanish, because *mecen-a-s* —whose diminutive is *mecenit-a-s*— is an *a*-stem but *Sócrat-e-s* and *Hércul-e-s* are *e*-stems, the same does not hold for Catalan. All these words are *e*-stems in Catalan (in Bermúdez-Otero's terminology), and thus there should be no difference in their behavior. Lloret (2015), who does not support the infixal analysis, considers that the words that retain the *-es* ending after the diminutive suffix *-et/-a* are interpreted as pseudofeminines. The parsing consists, thus, of a feminine marker /a/ plus a plural morph /z/: /mesén-a-z/ > /mesen-ét-a-z/. If, by contrast, they behave regularly in the diminutive production, as *Socratet*, it implies that the *-es* ending of the base is considered an inflectional marker as a whole: /sókrat-ez/ > /sokrat-ét/. Although this explanation is appealing because, if adapted into Catalan, it would also meet the subcategorization requirement of the infix *-et/-a*, as Spanish *-it-o/-a*, it is not clear to me why speakers should categorize words such as *mecenes* or *Sòcrates* in different ways. As opposed to Spanish, they all sound the same (as input for the hearer) and, thus, there should be no differences in their decomposition.[34]

---

[34] All forms being equal, the stress pattern in these nouns is different, as *Sòcrates* and *Hèrcules* are proparoxytones and *Mecenes* is a paroxytone. The same holds for Spanish, but I do not know

Alternatively, the differences in the diminutives of these forms might have to do with a language contact phenomenon. That is, it could be that this kind of infixation is not a genuine Catalan process but that the contact with Spanish has led to the segmentation of Catalan words with the same strategy used in Spanish. The asymmetries and the inconsistencies found between Catalan and Spanish are due to the superposition of the Spanish pattern to Catalan, which causes anomalous behaviors due to their different noun systems. In particular, the difference between *e*- and *a*-stems in words such as *Maties*, *Sòcrates*, etc. do not exist and speakers probably import the Spanish pattern. In other words, while the infixation subcategorization requirement works for Spanish, with a difference between *-o* and *-a* stems, on the one hand, and *-e* stems, on the other hand, it does not work for Catalan, where all nouns are conflated into the *-e* category. Nonetheless, they treat *mecenes* and *Sòcrates* as different due to the influence of Spanish. Speakers, also, may produce divergent outputs such as *Carlet* and *Carletes* for the same reason, as they are confused and do not know which procedure is the most appropiate for a word like *Carles* (something which would never happen in Spanish, where the form *\*Carlito* is completely ungrammatical).

The same explanation is applied to masculine words with an *-a* ending such as *drama* 'drama' or *mapa* 'map', which alternate, in their diminutive forms, between *drameta* and *dramet* 'little drama' or *mapeta* and *mapet* 'little map'.[35] As Lloret (2015) points out, the 'genuine' Catalan pattern does not extend the *-a* marker to the diminutive, yielding the regular masculine diminutives *mapet* and *dramet*; again, she attributes a pseudoinflectional interpretation to the *-a* ending (i.e., *-a* signals that the word is feminine). Significantly, Spanish only has *mapita* and *dramita* as diminutives and it can never generate outputs like *\*dramito* or *\*mapito* for *-a* masculine nouns. Under my analysis, in *mapeta* and *drameta* *-et-* works as an infix —(40), as in *mecenetes*— while in *mapet* and *dramet*, it works as a suffix (therefore [[map-a]–et-V] as in (41)):

(40)    map<**et**>–a

(41)    map–et

---

if prosody plays any role in the creation of the diminutives. In the *-us* and *-is* words analyzed previously in this section, there may also be a prosodic conditioning —in that case some kind of minimality effect— as words like *virus* and *brindis* have *viruset* and *brindiset* as diminutives but *eucaliptus* and *clitoris* have *eucaliptet* and *clitoret* instead (examples taken from Lloret 2015).

[35]Interestingly, words that have two genders, such as *artista* 'artist', can only generate the diminutive *artistet* for the masculine.

If Catalan speakers are hesitant between these two forms and the behavior is the same as in the case of the pseudoplurals (infixation vs. suffixation), the hypothesis about the influence of Spanish seems to be borne out, keeping in mind that the suffixation pattern is the most traditional in Catalan. Once the infixation process of Spanish has been imported into Catalan we can understand the variability found in masculine -*a* nouns and pseudoplurals.

Something similar happens in feminine words with an -*o* ending, which are extremely rare. Only *foto* 'photo' and *moto* 'motorcycle' are of regular use —and both are shortcuts for *motocicleta* and *fotografia*—, as well as *ràdio* 'radio'. According to what was just mentioned, we should find the variants *\*moteto* (infixation) and *moteta* (suffixation), but the diminutives *\*moteto* or *\*foteto* do simply not exist in Catalan (cf. *fotito* and *motito* in Spanish, which are very common). Instead, Lloret (2015) reports a form *motet*, considered there a pseudomasculine, assuming that speakers interpret -*o* as a masculine marker —there is no other possible explanation as why this noun does not bear the regular feminine inflectional ending -*a* in the diminutive.[36] However, this form is unacceptable for most speakers, which accept *moteta* as a better candidate, although this is not a very good candidate either. In fact, all these 'special' nouns prefer to use the adjective *petit/-a* 'little' + noun instead of creating a diminutive (*moto petita* is much more common than *moteta*). The fact that infixation is not an option here could invalidate the language contact analysis described above about the nature of infixation, but we have to keep in mind that there are only three words that follow this pattern and that speakers hardly ever use them, so they will be simply treated as exceptions.[37][38]
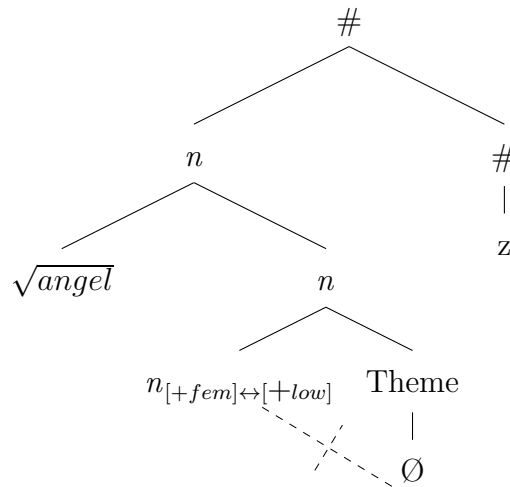
We still need to account for pseudoplurals with -*s* such as the feminine proper name *Àngels* 'angels' (derived actually form *Nostra Senyora dels Àngels* 'Our lady of the angels'), or the masculine common noun *temps* 'time'. In the case of *Àngels*, the proper name is clearly a pseudoplural because it is used for a singular (and feminine) referent, but its shape coincides with the masculine plural of *àngel* 'angel', *àngels*. Therefore, it needs to be specified as [+fem] in the lexical entry, but it has the same structure as the masculine common noun, as in (42):

---

[36]Lloret (2015, p. 72) argues that the suffix -*et/-a* cannot bear a masculine marked exponent, i.e, -*o*, and that "it only admits the inflectional regular endings : -Ø, -*a* and -*s*" (*només admet les terminacions regulars de flexió: -Ø, -a i -s*).

[37]These are xceptions because they cannot infixate and create diminutives such as *\*moteto* and *\*foteto*, but the outputs *foteta* and *moteta* are easily explained under the present analysis as they simply add a suffix to the noun.
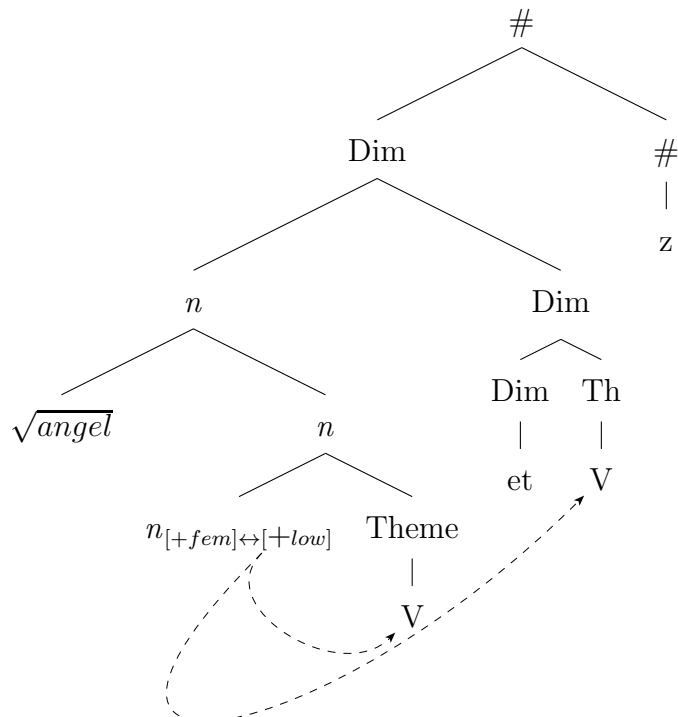
[38]Spanish tends to use more diminutives than Catalan —especially certain varieties, mostly in Latin America— but I do not know if this is related to the lack of some forms in certain contexts in Catalan.

(42)  *Àngels* 'angels'

```
                              #
                          /       \
                        n           #
                     /     \        |
               √angel       n       z
                          /    \
              n[+fem]↔[+low]    Theme
                         ⤫        |
                          ` ` ` ` Ø
```

Even though it behaves as a masculine noun, it is shown in (42) that *Àngels* bears feminine syntactic features, and therefore it needs to be stored with a -Ø ending to prevent the appearance of -*a* (if it were spelled out as a default noun, a V-slot would be supplied for the insertion of the inflectional ending). Besides this, we need to account for the two possible diminutive forms that are created from it, *Angeletes* i *Angelets*. In the first case, the speaker first produces the feminine diminutive of the word *Àngela* (also a feminine proper noun, derived from the common noun *àngel* 'angel' by the adjunction of the feminine marker -*a*), *Angeleta*, and then attaches the plural morph /z/, thus yielding *Angeletes*, shown below:

(43)  *Angeletes* 'little Angels'

```
                                              #
                                          /       \
                                        Dim         #
                                     /       \       |
                                   n           Dim    z
                                /     \       /    \
                          √angel       n    Dim    Th
                                     /    \    |     |
                        n[+fem]↔[+low]  Theme et    V
                                             |
                                             V
```

Under Lloret's analysis, both *Angeletes* and *mecenetes* would bear an *-a* ending, interpreted in *mecenetes* as a pseudofeminine, but in my proposal only *Angeletes*, a real feminine noun, selects it, while *mecenetes* bears a lexicalized theme vowel *-e* and the infix *-et-*.

*Angelets*, instead, comes from the plural form of the masculine diminutive form *angelet*, derived in turn from *àngel*, although it is specified as [+fem]. The two options, then, come from the same common noun *àngel*, but in *Angeletes* it is derived from a feminine form and in *Angelets* it comes directly from a masculine noun which is specified as feminine.

Things are more complicated with regard to *temps* 'time'. Some speakers consider *-s* to be part of the root, as the diminutive *tempset* shows, but others tend to create diminutives without the *-s* ending, as in *tempet*. Some derivatives, such as *temporal* 'temporary' or *temporada* 'season', do not retain the *-s*, and therefore two allomorphs ({temp, temps}) could be posited in the lexical entry of *temps* (the same that was assumed for *virus* 'virus' above). When confronted with the same word in Spanish, *tiempo* 'time' (which bears the default *-o* ending of masculine nouns), the irregularity of *temps* becomes clear. While forms such as *tempet* or *tempset* are extremely rare, the diminutives *tiempito* or *tiempecito* are very common in Spanish. This clearly indicates that the *-s* ending poses some problems for parsing as several decompositions can be inferred from the same input.

In fact, *temps* is special not only in derivative processes but also in its plural formation. As a masculine word whose singular contains a sibilant in final position, we would expect a plural *\*tempsos*, as in *gos > gossos* 'dogs, dogs', but the actual form is *temps* as well, i.e., it is invariable. *Temps* is not the only word that behaves in this particular way, as other nouns, both masculine and feminine, are invariable with regard to number, as shown in (44):

(44)

| | | |
|---|---|---|
| *pols* | (f.) | 'dust' |
| *calç* | (f.) | 'lime' |
| *tos* | (f.) | 'cough' |
| *bilis* | (f.) | 'bilis' |
| *índex* | (m.) | 'index' |
| *cactus* | (m.) | 'cactus' |
| *llapis* | (m.) | 'pencil' |
| *dúplex* | (m.) | 'duplex' |
| *fons* | (m.) | 'depths' |
| *dilluns* | (m.) | 'Monday' |

There have been several analyses regarding these nouns where *-s* is interpreted as a special morphological marker or an extraprosodic element (see Mascaró 1986; Wheeler 1979; Lloret 1995, among others), something that causes problems when the plural morph /z/ is adjoined to the base. However, none of them can explain the behavior of certain nouns with respect to similar cases that pattern regularly in a systematic way. Take, for example, the masculine noun *cens* 'census'. Its plural *censos* behaves as expected, with the use of the masculine morph /o/ as a means to avoid the contact of two sibilants. By contrast, the noun *ens* 'entity' (also masculine and with the same phonological shape except for the onset 'c'), with *ens* as a plural, is invariable. And the same holds for *fons* 'depths', that is invariable as well but has the same phonological structure as *cens*.

In other cases, however, it could be argued that the creation of plurals is prosodically conditioned. Regular plural formation in paroxytone nouns such as *virus* or *bilis* would yield outputs like \**v*[í]*rusos* or \**b*[í]*lises*, with stress on the antepenultimate syllable, something that the phonology of Catalan tries to avoid. If the stress shifted to the penultimate syllable, \**vir*[ú]*sos* or \**bil*[í]*ses*, there would be a violation of faithfulness with respect to the base, so fusing the *-s* to the stem seems an optimal solution. The same holds for a paroxytone word like *llapis*, but many speakers, although it is considered invariable in grammar books, produce the form *ll*[á]*pissos*.[39] In any case, this explanation cannot deal with oxytone words such as the already mentioned monosyllabic *fons* and *ens* or the disyllabic *dilluns* 'Monday' (or the other days of the week that have an *-s* ending, *dimarts* 'Tuesday', *dimecres* 'Wednesday, *dijous* 'Thursday' or *divendres* 'Divendres'). A plural \**dillunsos* would be 'perfect' form a morphological and phonological point of view, but it is not found in the language.

Alternatively, one can argue that these invariable words are all treated as pseudoplurals, which means that number is already spelled out in their morphological structure and thus no other plural morph can be added (see Harris 1992, 1999 for the same idea in Spanish). We saw, however, that in forms such as *virus*, with the diminutive *viruset*, the *-s* is part of the root, so this explanation cannot be applied to all cases.

Given the differences mentioned above and acknowledging the very few cases that deviate from the regular pattern, I conclude that the invariability of this set of nouns is purely idiosyncratic and must be learned. There is no (apparent) reason

---

[39]Interestingly, words ending in unstressed *-is* remain —with the exception of *llapis* for most speakers— invariable (*brindis*, *iris*, *clítoris*,...), whereas in Spanish both options are found, e.g., *iris* > *irises* but *brindis* > *brindis*. According to Bermúdez-Otero (2006b) this due to their categorization as *e*-stems or athematic.

why *cens* should be regular but *fons* not. As the list that conform this group is very limited, they must be stored in the lexicon as exceptions, specifying that the singular and the plural are the same.

Leaving aside plural formation, until now we have only tested the morphological structure of nouns with the diminutive suffix *-et/-a*. There are other expressive suffixes, like augmentatives, that have been left aside but they can also help us better understand the decomposition of Catalan nouns.

The most common and productive augmentative suffixes are *-ot/-a*, *-às/-sa*, *-arr-o/-a*, exemplified in (45), which roughly imply greater size or intensity (although they can have affectionate meanings as well):

(45)   *casa*   'house':   *cas-ot-a*   *cas-ass-a*   *cas-arr-a*

      *gitano* 'gipsy':  *gitan-às*   *gitan-ot*   *gitan-arr-o*

Augmentatives (almost) always behave regularly. That is, while in the diminutive there is variability between forms that retain exceptional word endings, such as *mapeta* 'little map' and forms that do not, such as *mapet* 'little map' (both from the masculine noun *mapa* 'map'), these exceptional markers never show up in augmentatives. Thus, forms like *\*mapassa*, *\*maparra* or *\*mapota* are totally unacceptable; only the expected masculine regular outputs *mapàs*, *mapot* and *maparro* are generated instead. This is a very systematic process and even problematic nouns with *-es* endings, such as *mecenes*, pattern as expected with the form *mecenot* or *mecenàs*, never *\*mecenotes* or *\*mecenasses* (keep in mind, though, that these forms are extremely infrequent). Lloret (2015, p. 73) provides *Matiotes* and *Eliotes*, from *Maties* and *Elies* (that actually have *Elietes* and *Matietes* as diminutives) as possible outputs, but most speakers would prefer *Eliot* and *Matiot* in the eventual context of producing such forms.

Bermúdez-Otero (2006b, 2007)'s arguments about the special infixal character of Spanish *-it/-o* (extended here to Catalan *-et/-a*) are powerful enough to support the analysis proposed in this section. That is, if endings such as *-(e)s* and *-a* (in the masculine, e.g., *mapa*) were interpreted as pseudofeminines (according to Lloret's analysis) we should also expect augmentatives with the 'false' endings, not *only* in diminutives. By contrast, if one assumes that it is precisely *-et/-a* what triggers the unexpected behavior, it follows that suffixation proceeds normally in augmentatives (it produces regular outputs). In other words, a pseudointerpretation should have consequences in all derivational environments, as opposed to a specification in *-et/-a* that restricts the infixation procedure to the diminutive. Both Lloret and Bermúdez-Otero's proposals have to acknowledge that *-et/-a* behaves in a particu-

lar way different from the augmentatives *-ot/-a*, *-às/-sa*, *-arro/-a*, but the present analysis (along the lines of Bermúdez-Otero) can account for the different outcomes of the diminutives without resorting to a) an *-es* inflectional pseudoplural marker and b) an *-a-s* ending consisting of a pseudofeminine plus a pseudoplural marker. By assuming that *-et/-a* can behave infixally, these differences vanish.

**7.1.1.2.3   Other 'athematic' endings**   Apart from the -(V)s endings analyzed above, there are other endings that pose some problems from a morphological point of view. I have grouped them as athematic due to their special character, but they behave as thematic in some respects.

Within this category, we find masculine and feminine nouns with *-u* and *-i* endings. Some masculine nouns have an *-i* or *-u* ending but must be treated as regular since these vowels are part of the root and found in derivatives. This is the case of words like *ritu* 'rite' and *divorci* 'divorce', whose structures, depicted in (46), are equivalent to those of *nen* 'kid' —(12):
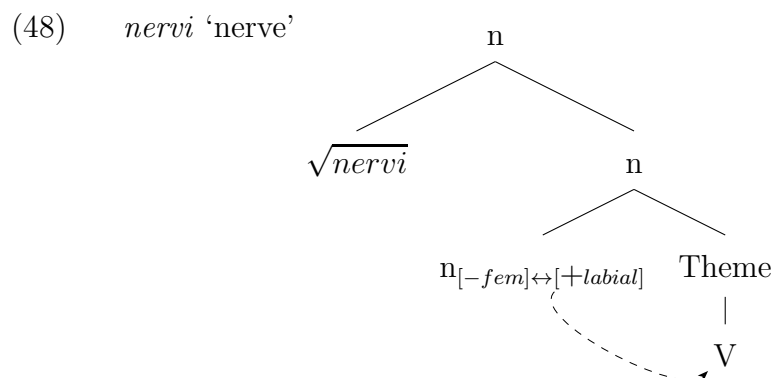
(46)    *divorci* 'divorce'
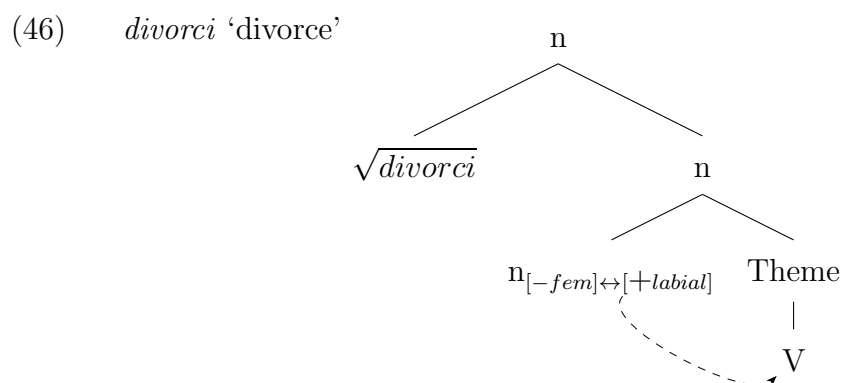


As a regular noun, it should get /o/ as its exponent, but, as argued in this work, in the phonological component the vocalic features that correspond to /o/ are not parsed. That *divorci* and *ritu* include *-i* and *-u*, respectively, in their roots is confirmed by derived forms such as *divorciar-se* 'to divorce' or *ritual* 'ritual'.

By contrast, there are other words in which there seems to be an inflectional ending *-i* or *-u* that does not show up in all contexts. Mascaró (1986, p. 97) provides some examples of this group, shown in (47):
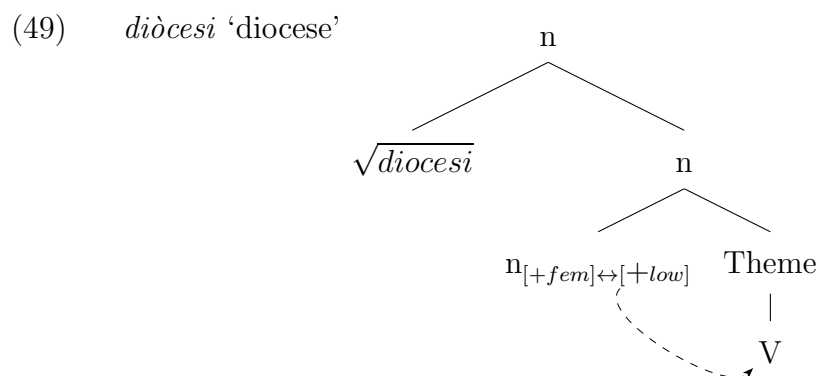
(47)    *nervi* 'nerve'  -  *nerviós* 'nervous'  -  *nervadura* 'nervation'

        *dosi* 'dose'  -  *dosificar* 'ration'  -  *dosar* 'to ration'

This set of nouns must include two allomorphs in their lexical representations, as we have already done for *virus* and *brindis* (cf. *viruset* 'litlle virus' but *víric* 'viral; brindiset* 'little toast' but *brindar* 'to toast'). Otherwise, it is impossible to account

for the differences they show in derivatives. The fact that they present this pattern of allomorphy is what motivates their classification as athematic, but it is clear that they get their inflectional exponents by regular means (which implies, in the masculine, not getting any vocalic features). Therefore, a word like *nervi* is represented exaclty like *divorci*, as in (48):

(48)    *nervi* 'nerve'

$$
\begin{array}{c}
n \\
\diagup \diagdown \\
\sqrt{nervi} \qquad n \\
\diagup \diagdown \\
n_{[-fem]\leftrightarrow[+labial]} \quad \text{Theme} \\
\mid \\
V
\end{array}
$$

The only difference between *divorci* and *nervi* is that, in the latter case, two allomorphs, {nervi,nerv}, are stored in the lexicon, using each one in different contexts, while in *divorci* there is only one for all contexts, {divorci}. In some cases, however, the *-i* marker never shows up in derivatives, as in *bigotet* 'small moustache', *bigotut* 'with a big moustache', etc. from the masculine noun *bigoti* 'moustache'. While we could consider this a 'real' inflectional ending, I think we should extend the idea of allomorphy here as well (along the lines of Viaplana 1991 and Lloret 2013), as speakers must categorize it with the other nouns that really alternate between two forms, which implies that *bigoti* should have the same structure as *nervi* in (48). It is worth mentioning that feminine nouns in *-i* or -u, like *diòcesi* 'diocese' and *tribu* 'tribe' must behave like other feminine nouns with no overt exponent such as *sal* 'salt, which have lexicalized structures where the theme position bears a zero morph, as in (49):

(49)    *diòcesi* 'diocese'

$$
\begin{array}{c}
n \\
\diagup \diagdown \\
\sqrt{diocesi} \qquad n \\
\diagup \diagdown \\
n_{[+fem]\leftrightarrow[+low]} \quad \text{Theme} \\
\mid \\
V
\end{array}
$$

Therefore, these nouns follow a regular inflectional process when they are masculine but not when they are feminine, because otherwise the forms *\*diòcesia* and *\*tribua*

would be found.

## 7.1.2 Adjectives

Adjectives, as part of the nominal system, behave generally like nouns. However, I have included them in a separate section because they display some special characteristics that need to be analyzed in detail.

The regular inflectional pattern of adjectives mirrors that of nouns, with four different forms that bear the default exponents for masculine and feminine plus the plural morph in every gender, as shown in 7.3 with the adjective *petit* 'small':

|  | **Singular** | **Plural** |
|---|---|---|
| **Masculine** | *petit* | *petits* |
| **Feminine** | *petita* | *petites* |

Table 7.3: *Regular inflection in Catalan adjectives*

Underlyingly, all the paradigm is derived from a single morphological template (the same as in nouns), shown in (50):

(50)     Root   +   Theme Position   (+   Plural)

Adjectives spell out the phonological features that correspond to [+fem] or [–fem] by the same process that was assumed for nouns. Thus, a feminine adjective like *petita*, depicted in (51), will have the same structure as any regular noun, although its head will be little $a$:[40]

(51)     *petita* 'small'
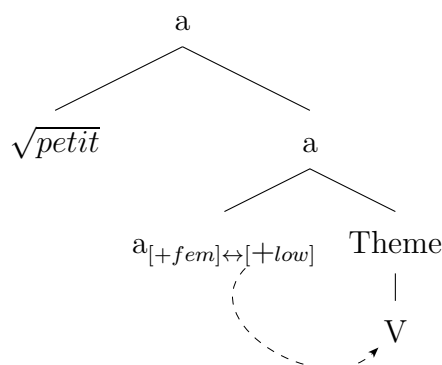


The same phonological constraints that controlled exponence in nouns are active for adjectives. The ranking PARSE[+LOW] >> DEPLINK guarantees the linking

---

[40]Adjective heads do not bear gender syntactic features, but I have placed them on little $a$ for descriptive purposes. Gender, as already mentioned, is assumed to be stored on little $n$, and adjectives inherit this feature from the noun they agree with.

of the vocalic features that conform /a/ in the feminine, while in the masculine
DEPLINK >> PARSE[+LABIAL] favors a candidate with no overt exponent (i.e., the
non-parsing of /o/).[41] In the plural, the adjunction of the morph /z/ to the singular
feminine and masculine adjectives creates the four cell paradigm shown in table 7.3.

Most adjectives, thus, have the four forms just mentioned. Leaving aside very
common phonological processes that apply to adjectives as well and generate 'dis-
cordant' masculine-feminine pairs, like word-final -n deletion (sa[Ø]-san[a] 'healthy',
underlyingly /san/-/san+a/) or cluster simplification (al[Ø]-alt[a], underlyingly /alt/-
/alt+a/), the main differences arise in adjectives that deviate from the regular pat-
tern (in the examples of n-deletion and cluster simplification above, the endings
are still the default), i.e., with secondary inflectional endings. The alternations are
basically the following (table adapted from Clua 2002, p. 507):

| Masc. | Fem | Examples | | |
|-------|-----|----------|---|---|
| o | a | *fond*[o] | *fond*[a] | 'deep' |
| e | a | *còmod*[e] | *còmod*[a] | 'comfortable' |
| a | a | *sibarit*[a] | *sibarit*[a] | 'sybarite' |
| e | e | *mediocr*[e] | *mediocr*[e] | 'mediocre' |
| w | a | *europe*[w] | *europe*[a] | 'european' |
| Ø | Ø | *difícil*[Ø] | *difícil*[Ø] | 'difficult' |

Table 7.4: *Non-regular inflectional endings in Catalan adjectives*

The pairs in table 7.4 show inflectional markers that were already analyzed in
the nouns' section, and hence no new observations (except for -w in the masculine,
analyzed below). In the case of masculine adjectives like *còmod**e***, *fond**o*** and *sibarit**a***,
their theme vowels are lexicalized as in their nominal counterparts (like *pare* 'father',
*lloro* 'parrot' or *atleta* 'athlete'), and thus they show the following structure:

---

[41]I am assuming that there is one single morph /a/ for the feminine, even in plural contexts
where its phonetic outcome is [e] —*petit*[e]*s*— (in Northwestern dialects like Pallarese, not in
Central Catalan where there is vowel reduction and in both singular and plural environments is
phonetically [ə]. Mascaró (1986, p. 96) argues that there is /a/∼/e/ allomorphy in the feminine,
but I follow Jiménez (2002), among others, in considering [e] a closed variant of /a/ when preceding
an inflectional morph.

(52)   *fondo* 'deep'
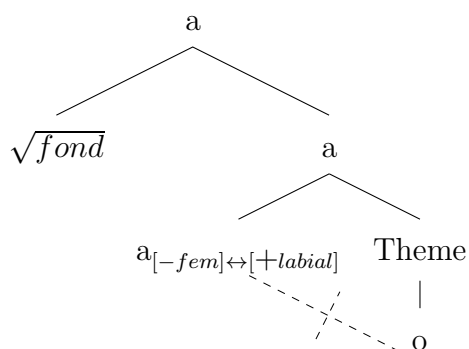
(53)   sibarita 'sybarite'
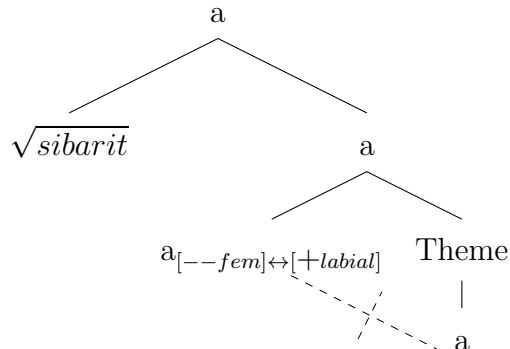
These lexicalized vowels block the application of the regular exponence process, which means that outputs like *\*fond \*còmod* or *\*sibarit*, with no overt exponent, are ruled out. The same is applied to *mediocre*, where in both the masculine and the feminine forms the vowel -*e* is stored in the theme position in the lexical entry, as in (54) and (55):

(54)   *mediocre* (masc.) 'mediocre'
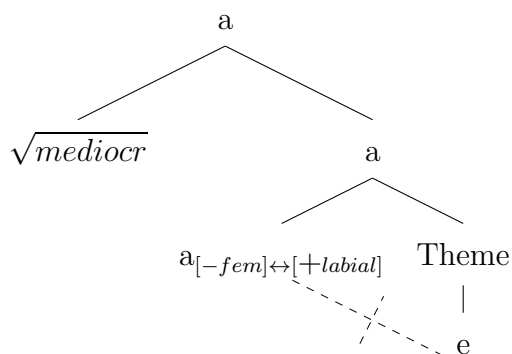
(55)   *mediocre* (fem.) 'mediocre'
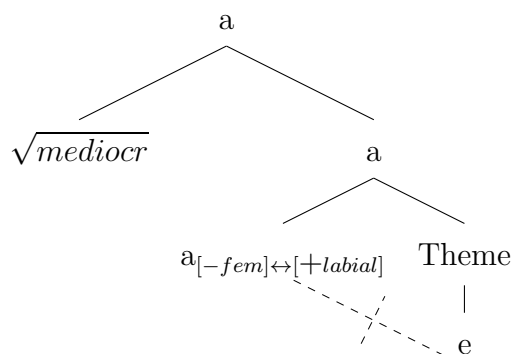
In traditional analyses, this vowel is considered epenthetic in the masculine (*\*medio**cr***) and lexical in the feminine, but I argued in the previous section that in masculine nouns like *centre* 'center' —which parallel the phonotactic situation of the word-final stop+liquid cluster in *mediocre*— the vowel must be lexical in nature as well because epenthesis is not active at the right edge of the word. If it were, the analysis defended in this thesis would predict forms like *\*centro* and *\*mediocro*, as the vowel /o/ remains floating and is only parsed in phonotactically illicit environments like this. The vowel in the feminine adjective *mediocre* must be considered lexical in any case (cf. *\*mediocra*), so extending the analysis to the masculine does not seem more costly to me.

So far, the morphological assumptions for nouns' structures were also applied to adjectives, which basically consider that any non-regular exponent is stored in

the lexical entry. There remains a group of adjectives, though, that have not been treated yet, which alternate between -[w] (masculine) and -[a] (feminine) or have directly no exponence at all in the feminine (neither in the masculine, but that is the default case):

| Masc. | Fem | Examples | | |
|---|---|---|---|---|
| w | a | *europe*[w] | *europe*[a] | 'european' |
| | | *bla*[w] | *blav*[a] | 'blue' |
| Ø | Ø | *difícil*[Ø] | *difícil*[Ø] | 'difficult' |
| | | *feliç*[Ø] | *feliç*[Ø] | 'happy' |

Within the -[w]/-[a] group, the -[w] marker displays a special behavior. In masculine adjectives like *europe*[w], the vowel *-u* is specified in the theme position of the lexical entry, as in (56).[42] Its feminine counterpart *europea* is instead a regular adjective, as shown in (57):

(56)   *europeu* 'european'          (57)   *europea* 'european'
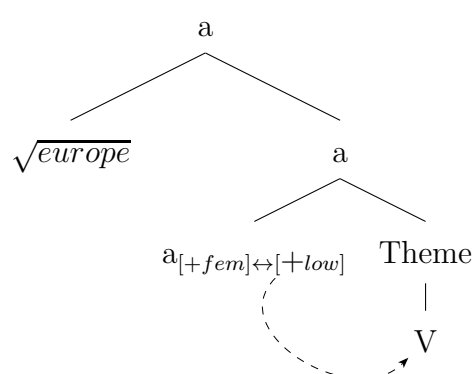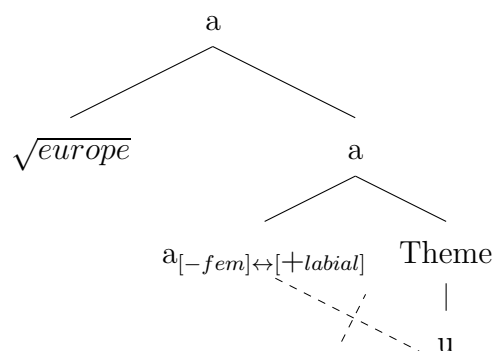


Again, the feminine inflectional ending will be realized in (57) due to the the phonological constraint ranking, yielding *europe*[a], while in the masculine, in (56), the features of *-o* will not be parsed (i.e., *\*europeuo*), as expected.

The behavior shown by diminutives like *europeuet*, which retain the labiovelar, as opposed to the feminine *europeeta*, goes against considering -[u] an independent morphological marker. Lloret (2013) argues that -[w] is retained in derivatives because it avoids a hiatus, although in *europeeta* there is also a hiatus. It is clear, though, that if we treat *-u* as a theme vowel specified only for the masculine it can

---

[42]Mascaró (1986, pp. 97-98) argues that [w] is underlyingly /u/, as words like *tio* 'uncle', [ti.u] in Central Catalan, show that it cannot be derived from /o/ because this vowel does not become a glide before another vowel, contrary to *europe*[w]. Mascaró considers as well that *-ew* and *-ea* are in fact *-e+u* and *-e+a*, being *-e* a derivational suffix —cf. *mar-e-a* 'tide' from *mar* 'sea'— but I will treat them as part of the root in order to simplify the exposition of the facts.

never be used in the feminine, so there is no other option than having a hiatus there. By contrast, in the masculine, we can improve syllable structure by using *-u*.

The ending *-*[w] can also appear in nouns (*dé*[w] 'god', *atene*[w] 'cultural center', etc.). The same analysis applies to them.

We have to carefully differentiate these cases from other adjectives that, on the surface, look almost the same. This is the case of *bla*[w]-*bla*[βa] 'blue' or *jue*[w]-*jue*[βa] 'jewish', with an alternation [w]-[β] (plus the regular feminine inflectional ending [a]) in masculine and feminine forms which is crucial to treat these forms as different from *europeu-europea*. On the surface, we have two different sounds, [w] and [β], but they can be derived from the same phoneme, /v/.[43] This implies that the root contains /v/ in both genders, being their underlying representations /blav/ (masculine) and /blav+a/ (feminine). Therefore, they are not equivalent to cases like *europeu-europea*, where the ending *-*[w] is only present in the masculine, never in the feminine. Their structures are regular, as shown in (58) and (59):

(58)    *blau* 'blue'                    (59)    *blava* 'blava'
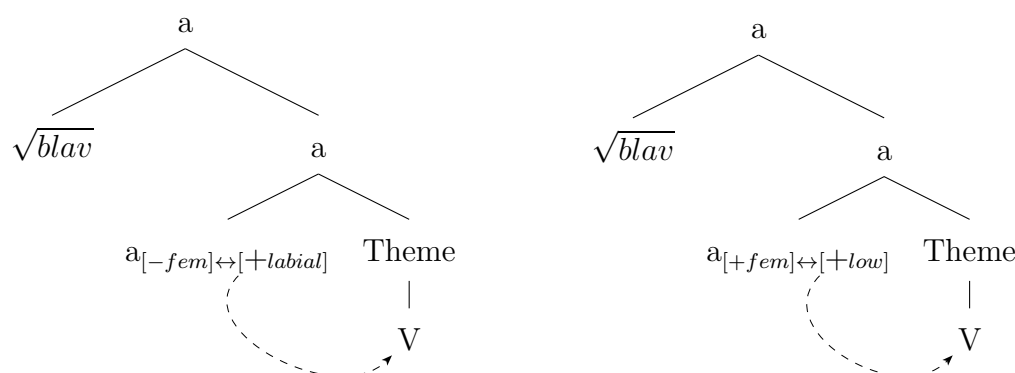


The phonetic differences between *blau* and *blava* are due to regular phonological processes of Catalan, but, as shown in (58) and (59), their morphological structures are equivalent to any regular masculine-feminine pair, like *nen-nena* 'kid'.

Feminine adjectives with a *-Ø* ending also pose some problems for the analysis. They can be divided into two groups, one where feminine and masculine have exactly the same shape in both the singular and the plural and one where the feminine and the masculine share their form in the singular but not in the plural. In other words, adjectives with two —(60)— or three —(61)— forms in the paradigm:

---

[43]Although this is irrelevant for the morphological analysis, I follow Wheeler (1979) in considering [w] and [β] the phonetic variants of /v/. Other authors like Clua (2002) consider that it is derived from /b/, but then there should be an alternation [p]-[β] due to final devoicing, as in *llo*[p]-*llo*[β]a 'wolf'. Mascaró (1976, pp. 97–99) discusses these alternations and posits also a 'b to glide' rule, although he acknowledges that there must be allomorphy in some cases.
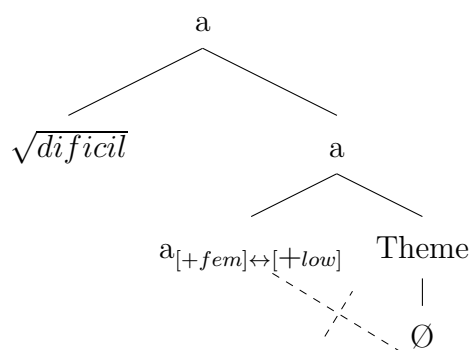
(60)     *difícil* 'difficult'          (61)     *feliç* 'happy'

|           | Singular | Plural |
|-----------|----------|--------|
| Feminine  | *difícil* | *difícils* |
| Masculine | *difícil* | *difícils* |

|           | Singular | Plural |
|-----------|----------|--------|
| Feminine  | *feliç* | *feliços* |
| Masculine | *feliç* | *felices* |

In both (60) and (61), the masculine singular forms behave as expected. They are regular and, thus, have no overt exponence. In the feminine, however, we have to assume, as we did with feminine nouns that have no overt exponent like *sal* 'salt', that these forms are lexicalized with a zero morph in the theme position, as in (62) and (63), which means that the regular process of inflection cannot apply and the outputs *\*difícila* and *\*feliça* are not generated:

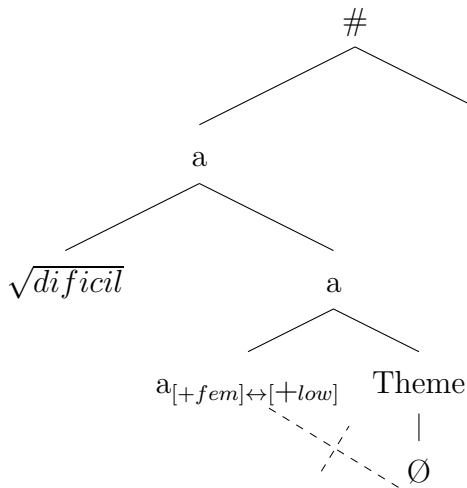(62)     *difícil* 'difficult'          (63)     *feliç* 'happy'
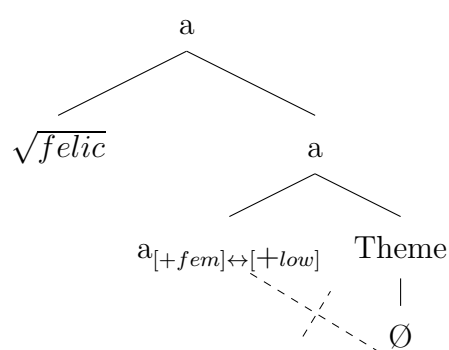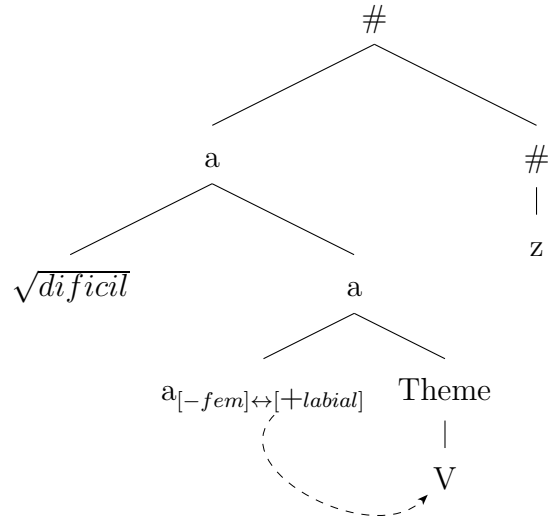


In the case of *difícil*, the plural does not pose any problem either, as it just attaches the plural morph /z/, yielding *difícils* in both genders, even if their morphological structures are different (lexicalized -Ø exponence in the feminine —(64)— vs. non-parsing of features in the masculine —(65)—):

(64)    *difícils* (fem.)                    (65)    *difícils* (masc.)



The difference lies in the theme position, as in the masculine it consists of a V-slot that will end up unpronounced because it lacks the place features corresponding to /o/ (which are not parsed due to the constraint ranking), while in the feminine there is no V-slot available (the structure bears a zero morph in the theme position, and thus no features at all attached to it).

These feminine structures are, in principle, lexicalized and do not bear the feminine marker /a/, but for some speakers there is overapplication of the default inflectional process in order to create a four-cell paradigm, and therefore they treat them like regular adjectives. It is not uncommon, then, to find pairs like *gegant-geganta* 'enormous' (*una casa geganta* 'an enormous house'), even when this adjective is usually invariable (*una casa gegant*) and has only one form in the singular and one form in the plural for both genders (*gegant-gegants*).

More problematic are the cases that are invariable in the singular but not in the plural, as in the paradigm of the adjective *feliç* 'happy' shown in (61). While the structures of *difícil* and *feliç* in (62) and (63), respectively, are the same and we must assume that they are simply lexicalized, the appearance of the vowel *-e* in the feminine plural poses some problems for the analysis that are detailed below.

In both the masculine and feminine plurals there is a phonotactic problem due to the contact of two sibilants, one belonging to the root and one belonging to the plural morph, /felis+z/. We saw in chapter 5 and also in the section dedicated to nouns of this chapter that, in the masculine, this OCP contact (*feliss*) is avoided by linking [+labial] into the V-slot of the theme position, as in (66):

(66)    *feliços* 'happy'
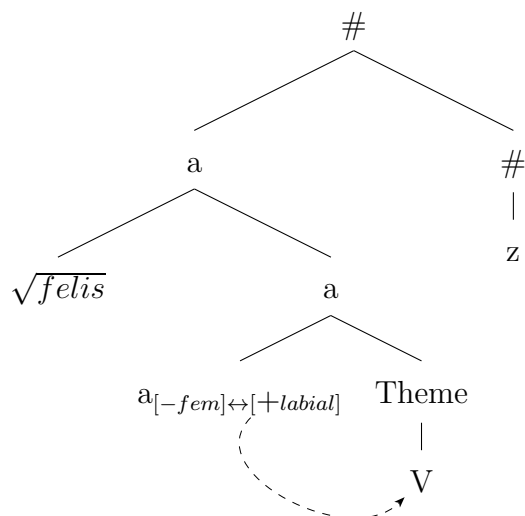


This is exactly the same process that applies to plurals such as *gossos* 'dogs' (from *gos*) or the phonologically conditioned alternation between the 3rd person masculine singular clitic forms *l/lo*. The constraint ranking licenses the linking of the features associated with masculine only when forced by phonotactics, and thus we can consider this a morphological rescue strategy. However, we cannot extend this analysis to the feminine plural adjective. In the masculine singular *feliç* the [+labial] features remain floating and are not pronounced after Stray Erasure. When creating the plural, though, the features are still floating, something which is not possible in the feminine because, as we saw previously, the singular adjective is stored with a zero morph. If the theme position is -Ø, the general inflectional process that would assign /a/ as the default exponent is blocked in both the singular and plural forms. Thus, no floating features are available to be inserted in the theme position when required by syllabic structure.

These kinds of adjectives constitute a challenge for the analysis. The presence of *-o* in the masculine plural in the theme position clearly shows that we are dealing with morphological structure (even if it is constrained by phonotactics), so at first glance it appears that the vowel in the feminine plural is a morphological marker as well. However, we are forced to assume that there is a zero morph in the feminine singular (there is no other way of blocking the default insertion of /a/), which should, in principle, also be maintained in the plural. How can we solve this puzzle, then? The idea is simple: if there is already a zero morph in the theme position (in both the singular and plural), there is no morphology-phonology mismatch if a vowel is epenthesized because it lies outside a morphological position. I conclude, then, as has been usually assumed in the literature (see Mascaró 1986; Clua 2002; Bonet et al. 2007; Lloret 2013 among others), that the vowel in *felices* is epenthetic, which implies that their morphological and phonological domains are the following (italics

indicate that -*e* is epenthetic):

(67)     ([[felis-Ø]*ez*])

In (67) it is shown how the vowel -*e* is outside the stem [felis-Ø]. The zero morph is occupying the theme position and, as a consequence, nothing else can be inserted there. As opposed to the plural clitic *les* (alternating with *ls* depending on the phonological context), there is no V-slot where [–back] features can be linked (precisely because the rule Th ↔ V is blocked by the zero morph), and thus the only possible solution is to epenthesize a whole vowel, i.e., vocalic features, association lines and an X-slot. Thus, the -*e* in *felices* cannot be considered a theme vowel.

It has been argued throughout this dissertation that epenthesis cannot apply at the right edge of the word, and the analysis of *felices* seems to contradict this idea. The difference with regard to other epenthetic environments is that in this case there is no other option than inserting a vowel. That is, epenthesis is a last resort strategy but, nevertheless, the only one. In a word like *centre* 'center', we conluded that the vowel is lexical and epenthesis should be ruled out because, as indicated by loanwords, we would expect a form *\*center* —with internal epenthesis— to avoid interferences with the theme positon (the word-final coda cluster can be repaired by two different means: /tɾ/ > [tɾ*e*]/[t*e*r]). In *felices*, though, the two sibilants create an OCP problem that can be repaired only by inserting a vowel between the root sibilant and the plural morph. Nothing else will solve the problem, and morphology can only help in the masculine plural, as there is no access to the default feminine exponent by the blocking effect of the zero morph. Lloret (2013, pp. 265-266) arrives to the same conclusion, as she argues, in the lines of Bonet et al. (2007), that lexical entries contain hierarchically ordered inflectional allomorphs (finally determined by the constraint ranking), but nevertheless their presence is blocked in *felices* by a diacritic in the root that selects -Ø (a similar argumentation to the one advocated here). Bonet et al. (2007, p. 924), instead, posit a constraint ranking that favors the candidate with epenthesis, but leave open the possibility to an inflectional interpretation of -*e*:

> There is another possibility, however: if RESPECT and ALIGN-MM are unordered or have the opposite ranking (that is, ALIGN-MM >> RESPECT), the candidate with the gender allomorph will be the optimal one, the candidate with epenthesis being ruled out by PRIORITY

An alternative to this analysis could be the storage in the lexical entry of multiple inflectional endings. This is what Bonet (2006) and Bermúdez-Otero (2013)
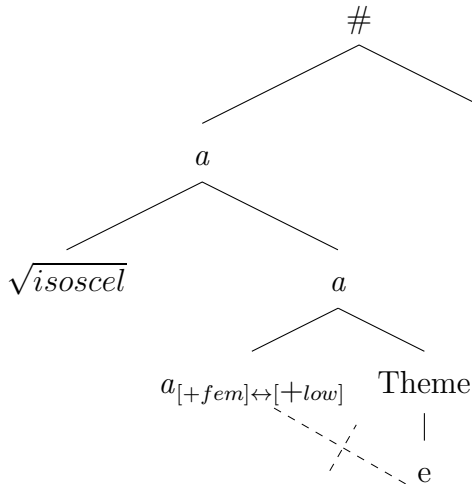
propose for Spanish *e*-nouns like *pan* 'bread', that have a plural *panes*. As there is a difference between these set of nouns and others like *fan* 'fan', which pluralize as *fans*, Bonet and Bermúdez-Otero assume that the vowel *-e* of *panes* 'bread loaves' is actually an inflectional ending that signals its belonging to the thematic class —as opposed to *fan*, which is athematic— and thus its lexical representation must be /pan$_{\emptyset,e}$/ or /pan$\{\emptyset,e\}$/ (Bonet's is a root-based approach while Bermúdez-Otero's is stem-based). Something similar could be posited for *felices*, with a lexical entry /felis$\{\emptyset,a\}$/ that would guarantee the presence of the zero morph in the singular and the feminine marker /a/ in the plural. While this is a technical solution that works for the cases studied here, it loses the phonological generalization that applies to this kind of adjectives. It is the OCP problem which motivates the presence of the vowel because it is otherwise absent, as in the feminine pair analyzed above *difícil-difícils*, and thus it would be just a coincidence that the adjective *feliç* selects /a/ in the feminine plural (something that could happen in *difícil* as well, which implies that there is no contrast between *difícil* and *feliç*).[44]

Moreover, Clua (2002, p. 511) argues that treating the vowel *-e* in *felices* as a gender marker contradicts universal tendencies by which singular elements must have gender distinctions if their plural counterparts have them as well. This is actually Greenberg's universal 37 ("a language never has more gender categories in nonsingular numbers than in the singular"), and, as we saw in chapter 5, this is borne out by the behavior of Pallarese 3rd person clitics, where gender distinctions are lost in the plural but maintained in the singular. The epenthetic analysis of *felices* adjusts to this tendency.

Finally, we need to treat the adjectives that are completely invariable, in both gender and number. We can group them with the pseudoplural nouns of section 7.1.1, and therefore their shape, reproduced in (68) (fem.) and (69) (masc.), is easily explained in both numbers:

---

[44]This is something similar to Paster (2006)'s analysis of phonologically conditioned allomorphy, which, under her view, should be reduced to subcategorization requirements (subcategorization is the reason why some allomorphs appear in certain phonological environments, but there is no optimization at all).

(68)  *isosceles* 'isosceles'     (69)  *isosceles* 'isosceles'
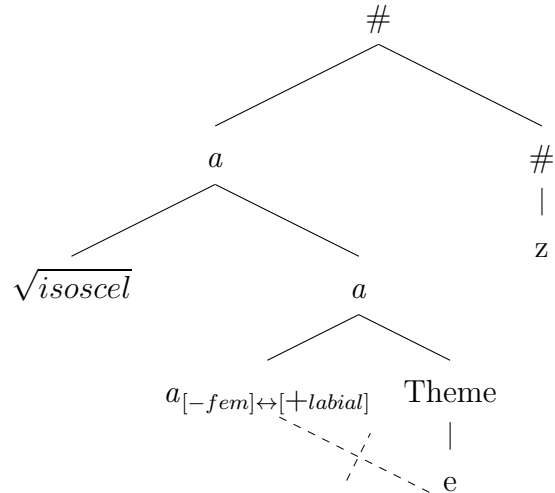


The theme vowel is occupied by *-e* in both masculine and feminine and there is a plural projection even if the adjective is singular for agreement purposes (*un triangle isòsceles* 'an isosceles triangle'). As the plural morph is already spelled-out, there can be no differences when the adjective becomes fully plural (*uns triangle**s** isòscele**s***).

### 7.1.2.1  Demonstratives

The distribution of demonstratives in Pallarese Catalan is the following:[45]

|  | Singular | Plural | |
|---|---|---|---|
| **Feminine** | aguesta | aguestes | 'this' |
| | aguella | aguelles | 'that' |
| **Masculine** | aguest | **aguests/aguestos** | 'this' |
| | aguell | aguells | 'that' |

Table 7.5: *Pallarese Catalan demonstratives*

In the feminine, demonstratives pattern as expected, i.e., with the default feminine ending *-a* in both the singular and the plural (as already mentioned, -[e]*s* in the feminine plural is derived from /a+z/). Some examples are shown in (70):

---

[45]In Central Catalan, the demonstrative is realized with a voiceless velar plosive, e.g. *a*[k]*est* 'this', whereas in Pallarese Catalan we find its voiced counterpart —an approximant due to inter-vocalic spirantization—, e.g., *a*[ɣ]*est* 'this'. Due to the influence of Standard Catalan, speakers use sometimes the voiceless plosive, but the use of the variant that contains the approximant is much more widespread.

(70)   a.  **[aɣésta]** iclésia   sembla  preciosa
            this.FEM.S church.FEM.S seem.3.S beautiful.FEM.S
            'This church seems beautiful'

       b.  **[aɣéstes]** tomates,     menja-    te'     ls!
            this.FEM.PL tomato.FEM.PL eat.IMPV.2S REFL.2S ACC.PL
            'These tomatoes, eat them!'

       c.  **[aɣéʎa]**   casa       d' allí
            that.FEM.S house.FEM.S of there
            'That house there'

       d.  **[aɣéʎes]** tomates,    amaniu-   mo'    ls!
            that.FEM.PL tomato.FEM.PL dress.IMPV.2S DAT.1PL ACC.PL
            'Those tomatoes, dress them for us!'

Gender inflection does not pose any problems for interpretation in feminine demonstratives, as they follow the regular process of phonological spell-out that has been developed throughout this thesis. The same holds for the masculine distal demonstrative, shown in (71):

(71)   a.  **[aɣéʎ]**    senyor
            that.MASC.S man.MASC.S
            'That man'

       b.  **[aɣéʎʃ]**   peixos     no  m'   agraden
            that.MASC.PL fish.MASC.PL NEG DAT.1S like.3PL
            'I don't like those fish'

Both in the singular and the plural, masculine is phonetically realized as -Ø, as the general process of gender exponence does not link the features associated with [–fem] in default cases. Therefore, it behaves as expected. In the proximal demonstrative, by contrast, some phonotactic constraints condition the appearance of -*o* in the plural, as indicated in table 7.5. In (72) the two surface forms for the plural proximal demonstrative are shown:

(72)   a.  **[aɣéts]**   cent   anys
            this.MASC.PL hundred year.MASC.PL
            'These hundred years'

       b.  **[aɣéstos]** peixos    no  m'   agraden
            this.MASC.PL fish.MASC.PL NEG DAT.1S like.3PL
            'I don't like these fish'

The demonstrative is underlyingly /agest+z/, but the complex coda -*sts*, which is in principle tolerated but almost absent in casual speech, is either simplified by deleting the first sibilant, thus yielding [aɣéts], as in (72-a), or repaired by linking the

masculine featural exponents (i.e., -*o*) to the theme position, as in (72-b).[46] This is in fact the same repair strategy used in Pallarese for the alternation in definite articles and pronomoinal clitics or in OCP-sibilant cases like *gos-gossos* 'dogs-dogs' or *peix-peixos* 'fish' (also in (72-b)). These cases have already been discussed in this and other chapters troughout the dissertation, and the same applies to the plural demonstratives. Furthermore, some theoretical models that focus on phonological variation, like Stochastic OT (Boersma, 1997), could explain the alternation [aɣéts]/[aɣéstos], but modeling variation is beyond the scope of this dissertation.

In the singular of the masculine proximal demonstrative, which is underlyingly /agést/, there is also consonant cluster simplification, and the output is generally [aɣét], as in (73):

(73)    **[aɣét]**    peix    no    m'    agrada
        this.MASC.S fish.MASC.S NEG DAT.1S like.3S
        'I don't like this fish'

Only when the demonstrative is followed by a vowel-initial word, the second element of the coda is syllabified as an onset and -*s* surfaces in the output ([a.ɣès.tɔ́.me]), as in (74):

(74)    **[aɣést]**    home    és    molt alt
        this.MASC.S man.MASC.S be.3S very tall.MASC.S
        'This man is very tall'

Therefore, the demonstrative surfaces faithfully only when followed by a vowel. According to the featural affixation analysis of Pallarese Catalan, a possible solution for the coda cluster -*st* before consonants should be the association of the masculine exponent to the theme position, that is, \**aguesto* (\**aguesto peix* \*[a.ɣès.to.péʃ], cf. (73)), but this is in fact never found in Pallarese. Only -*s* deletion is possible in the singular, whereas in the plural both options ara available: -*s* deletion or linking the features associated with [–fem].[47]

It is worth mentioning that, by paradigmatic pressure, also the forms [aɣéta] and [aɣétes] are found in the feminine, although there is no phonological motivation for -*s* deletion in the feminine proximal demonstrative. In Central Catalan, by contrast, -*s* is deleted in the whole masculine paradigm (it is found in prevocalic contexts in

---

[46]Some speakers may force the pronunciation [akésts] (Standard Catalan)/[aɣésts] due to stylistic reasons, but this is never found in spontaneous speech.

[47]This is an interesting asymmetry. As pointed out by Bonet (p.c.), the fact that the linking of -*o* is available in the plural may have to do with the presence of the plural itself. That is, the compulsory presence of the inflectional plural morph may license the use of another inflectional element, -*o*.

the singular, though, as in (74)), i.e., [əkɛ́t]/[əkɛ́ts], but it always surfaces in the feminine, i.e., forms like *[əkɛ́tə]/*[əkɛ́təs] are not possible, only [əkɛ́stə]/[əkɛ́stəs].

All in all, the behavior of Pallarese demonstratives shows once again that gender exponents are used as a repair strategy in case of phonological conflict.

## 7.2 Non-inflectional elements: adverbs

Adverbs pose a challenge for the analysis of Pallarese presented so far. They are clearly genderless elements, but nevertheless the well-formedness requirement posited by Oltra-Massuet (1999) affects them, and thus they project a theme position as nouns and adjectives. As adverbs do not bear gender features, the phonological spell-out proposal for gender developed in this thesis cannot target them, although sometimes they pattern together with nouns and adjectives in some respects, as shown in (75) (examples taken from Lloret 2015, p. 76):[48]

(75)  *lluny*   'far'                    >   *lluny-et*      'a little bit far'
      *enfora*  'towards outside'  >   *enfor-et-a*   'a little bit towards outside'

Adverbs can get diminutive suffixes, as the addition of *-et* in (75) shows, although it is not as common —by any means— to find them in adverbs as in nouns and adjectives (speakers vary in their degree of acceptance of these forms, too).[49] An analysis that posits morphological class markers (Harris 1985 and subsequent work) can easily explain the data in (75). As the relation between gender and word endings is not direct, adverbs can still bear class features and, thus, get a specific exponent for them (even if they are genderless elements, as opposed to nouns and adjectives). Therefore, an adverb can be stored with its corresponding class features and, after the adjunction of the diminutive suffix to the adverbial root, these features end up realized at the rightmost edge of the word. Recall, however, that the present analysis does not resort to class markers, as theme vowels are considered, instead, to be directly stored in lexical entries together with roots in nouns and adjectives which are not fully regular. The same idea can be applied to adverbs, thus yielding

---

[48]No diminutive forms for adverbs were found in the Pallarese data collected during the interviews. However, the morphological analysis developed in this section is valid also for other varieties of Catalan, and thus the Central Catalan examples of this section, taken from Lloret (2015) (who also provides some examples from Majorcan and Valencian Catalan), fit into the proposal.

[49]Lloret (2015) also reports some diminutive forms for other categories such as prepositions (only in intransitive prepositions which have an adverbial function like *a prop de* 'close to'), pronouns, demonstratives, numerals or interjections. As these are very marginal and non-productive cases, they are not considered here.

structures like those of (76):[50]

(76)    *enfora* 'towards outside'[51]
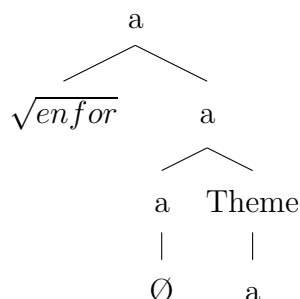


For an adverb like *enfora* in (76), the theme position is already spelled out and it is specified as -[a], i.e., it is lexicalized, as in non-regular nouns or adjectives (like *mar*[e] 'mother', for example). This vowel has no connection to gender whatsoever although it coincides with the phonological manifestation of [+fem]. The only thing in common that the feminine default exponent, -[a], and the vowel -[a] in *enfora* have is that they both occupy the theme position, which is pervasive in all syntactic functional heads, like little *n*, little *a* and so on. Thus, it is clear that both vowels are different in nature.

An adverb like *lluny* in (77), instead, does not have to specify any vowel in its lexical entry because it consists of a bare root. Nevertheless, the general process that spells out the theme position with an underspecified vowel (i.e., Th ↔ V) applies in adverbs as well, as this step is crucially dissociated from gender:

(77)    *lluny* 'far'
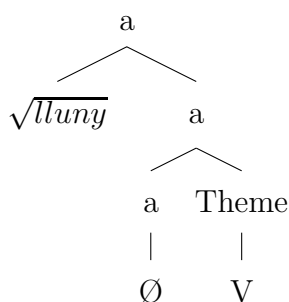


As there are no gender morphosyntactic values in adverbs, the place features associated with them can never end up in the V-slot of the theme position, and thus the [–cons] feature that conforms this V-slot is deleted after Stray Erasure, yielding the phonetic output [ʎúɲ].[52]

---

[50]Following Oltra-Massuet & Arregi (2005, p. 67), I consider adverbs to be of the same category as adjectives, thus containing the functional category *a*, but as noted by the authors, this assumption is not crucial, as they could bear an *adv* head.

[51]The adverb *enfora* is a combination of the preposition *en* 'in' and the adverb *fora* 'outside', which indicates directionality. It is treated in (76) as a simple lexical item for descriptive purposes.

[52]The place features of the vowel -*e* could be epenthesized in the V-slot of the theme position if there were a phonotactic problem, but as the palatal nasal is a licit coda in Catalan, there is no

The problem with an analysis of this kind (i.e., with lexicalized theme vowels) is that the use of diminutive suffixes in adverbs does not seem straightforward at first sight. That is, it has been shown that in proposals that resort to class features, adverbs bear those class features attached to roots, and therefore suffixes can be joined to them to create the diminutive and still spell out the exponents associated with class after the diminutive suffix. If, as I am arguing, there are no class features and the theme position is already occupied by a lexicalized theme vowel, how can we derive diminutives like *enforeta* 'a little bit towards outside' in (75)? The answer is, again (as already argued for some nouns in 7.1.1), that *-et* is not a suffix but an infix (see Bermúdez-Otero 2006b, 2007 for the same idea applied to Spanish), as assuming that the final vowel in *enforeta* belongs to a suffix would yield a structure like (78) (to be rejected, see (79) below):

(78)     [[enfor-**a**]–et-**a**]

In (78) it is shown that the theme positions of both the adjective and the diminutive heads are spelled out as -[a]. As gender is not part of the morphosyntactic composition of adverbs, it is clear that if there is any vowel occupying the theme position of the diminutive head it cannot be spelling out gender features and must be instead lexically specified.[53] Therefore, the move taken by some nouns or adjectives, whose theme position is only realized with the default vowel of the corresponding gender in the diminutive, cannot be extended to adverbs. That is, a feminine adjective like *gran* 'big' does not bear the regular ending *-a*, but this surfaces instead when a diminutive is created from the base, as in *grandet**a*** 'fairly big'. This can happen because, previously, the theme position of the diminutive head has been spelled out with an underspecified vowel so that the pleace features of [+fem] can be realized there. In adverbs this is not possible as there is no access to gender, and we should thus posit a lexicalized sufix *-eta* specifically for adverbs of this kind. As diminutive creation is not productive in adverbs, it seems more plausible to reject the structure in (78), where *-et* is a suffix, and consider *-et-* an infix instead, as in (79):

(79)     enfor<**et**>–a

The structure in (79) is the same that has been argued for in cases like *mapeta* 'little map' in (40) (which have been confronted with regular suffixal patterns like *mapet* 'little map' in (41) —for speakers that reject the infixation process). While

---

need to do so and, therefore, Stray Erasure applies.

[53]In a scenario like that of (78), the adverb would not surface as \**enfor**a**et**a***, with two realized theme vowels, by the rule that deletes the theme vowel of the inner head in derivational contexts that has been posited in 4.1.2 and 7.1.1.

infixation seems to be a synchronic process for some cases (and some speakers) in Catalan, it has been previously mentioned that this might be an influence of Spanish, which uses it more frequently. It is probably not a coincidence that adverbs with a -Ø ending (absence of phonological features) are preferred in their diminutive forms, e.g., *lluny > llunyet*. In adverbs that have a fully specified vowel in the theme position, like *enfor**a*** 'towards outside', speakers avoid the use of the diminutive precisely due to the vowel, which acts as a block to the process. When there is no such vowel, the attachment of the infix to the root is less problematic (nothing is lexicalized and therefore it ends up taking the shape of a suffix, because it appears at the rightmost edge of the word).

We can conclude, then, that the infixation analysis proposed for adverbs (as well as for some nouns, see section 7.1.1) is advantageous for two reasons:

- It explains why the creation of diminutives is not productive in adverbs.

- It explains the better acceptance of diminutives from bases like *lluny*, with no lexicalized theme vowel, with respect to bases like *enfora*, with a lexicalized theme vowel.

# CONCLUSIONS AND FURTHER RESEARCH

In this final chapter I briefly summarize the contributions of this thesis from both an empirical and theoretical point of view. In addition, I provide a cursory analysis of epenthesis in Central Catalan clitics in 8.2 that challenges the theoretical proposal defended in previous chapters for Pallarese, outlining possible directions for further research.

## 8.1  Proposals and Contributions

The contributions of this dissertation are twofold. Firstly, it provides new dialectal data about the distribution of clitics in Pallarese Catalan. To my knowledge, the data presented in chapter 2 constitute the most comprehensive description of the clitic system of the Pallarese variety spoken in Pallars Sobirà (specifically, in Vall d'Àneu and Vall Ferrera), and is a complement to the previous literature on the morphophonology of Pallarese (e.g., Coromines 1936, Sistac 1998, Viaplana et al. 2007, Alcover & Moll 2011 and Massanell 2012).

Secondly, this thesis contributes to a better understanding of the phonology-morphology interface and the mechanisms that control lexical insertion. Departing from previous analyses of Catalan and Spanish nominal inflection (reviewed in chapter 3.1), in chapter 4 I proposed a spell-out process for Pallarese Catalan gender in two steps, as illustrated in (1):

(1)     a.     Underspecified vowel in the theme position:

        Th ↔ V (*à la* Roca 2010)

        b.     Association of gender values with phonological features:

        [+fem]   ↔   [+low]            (= /a/)

$$[-\text{fem}] \quad \leftrightarrow \quad [+\text{labial}, -\text{high}] \quad (= /\text{o}/)$$

Given the well-formedness requirement that imposes a theme position on syntactic functional heads (Oltra-Massuet, 1999), an underspecified vowel (with no place features) is first spelled out in the theme position —(1-a)— of clitics or nouns and, subsequently, the floating features associated with gender —(1-b)— end up linked to the V-slot provided by this theme position, as in (2):

(2)    *casa* 'house'
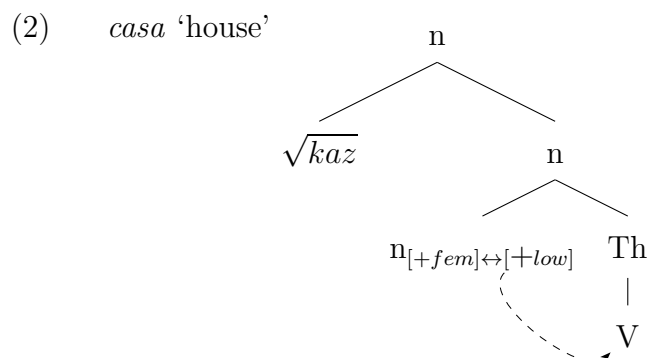


After this process, the feminine noun *casa* is fully realized as in (3) below:

(3)    *casa* 'house'
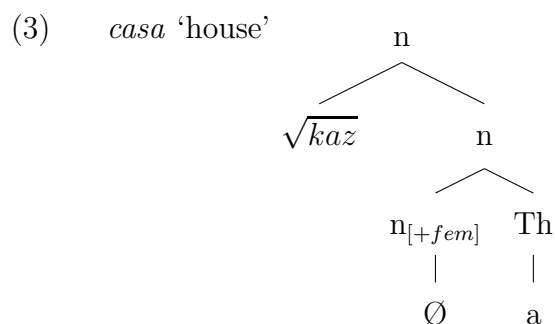


As illustrated in (1-b), the feminine and masculine exponents are -*a* and -*o*, respectively. Whereas -*a* is the most common inflectional ending for the feminine, this is not the case for -*o* in the masculine, with -Ø being the default pattern instead (*noi*[Ø] 'boy' vs. *noi*[a] 'girl'). The lack of realization of -*o* is considered to be derived from the constraint ranking shown in (4), which explains the different exponence between the feminine, with an overt inflectional marker, and the masculine, with an absence of phonological features on the surface:

(4)    PARSE[+LOW] >> DEPLINK >> PARSE[+LABIAL]

DEPLINK (Morén, 2001) goes against the insertion of new association lines in the output. Since the features associated with gender are floating in the input, any candidate that links those features to the theme position will violate DEPLINK. PARSE[+LABIAL] and PARSE[+LOW], by contrast, demand parsing of [+labial] and

[+low] features —the exponents for masculine and feminine, respectively— which implies linking them to the V-slot of the theme position in order to receive phonetic interpretation. Therefore, any candidate that satisfies DEPLINK will violate PARSE[+LABIAL]/PARSE[+LOW] and, conversely, any candidate that satisfies PARSE[+LABIAL]/PARSE[+LOW] will violate DEPLINK. DEPLINK dominates PARSE[+LABIAL], but DEPLINK is in turn dominated by PARSE[+LOW], thus explaining the differences in feminine/masculine exponence.

However, syllabification may force the surface realization of the features associated with masculine. In chapter 3, this has been shown for the 3rd person singular masculine accusative clitic (and definite article), which has two phonetic variants, [l] and [lo], as opposed to its feminine counterpart, with only one variant in the same phonological contexts (the underlying form for the clitic is /l–V/ in both the masculine and feminine). This asymmetry is illustrated in (5) and (6) below:

(5)  a.  No  **[l]**  porto  (6)  a.  No  **[la]**  porto
         NEG ACC.3S.MASC bring.1S        NEG ACC.3S.FEM bring.1S
         'I don't bring it'             'I don't bring it'

     b.  **[lo]**  porto        b.  **[la]**  porto
         ACC.3S.MASC bring.1S       ACC.3S.FEM bring.1S
         'I bring it'              'I bring it'

The dissociation of floating gender exponents and theme position in the inflectional system of Pallarese explains the use of *-o* in cases like (5-b) easily. The constraint ranking in (4) determines that, in the feminine, the floating features associated with *-a* are linked to the theme position both in (6-a) and (6-b), whereas in the masculine the ordering DEPLINK >> PARSE[+LABIAL] only favors a non-vocalic clitic in (5-a), where the clitic can be attached to the previous vocalic host. In (5-b), by contrast, the Sonority hierarchy (formally expressed in the constraint σSTRUC that dominates DEPLINK) bans an output with a liquid + plosive onset cluster, thus forcing the association of [+labial] to the theme position to satisfy syllabification requirements.

This use of inflectional vowels as a means to repair syllable structure has been labeled 'morphological epenthesis' in the literature (Cardinaletti & Repetti 2008; *epèntesi morfològica* in Lloret & Viaplana 1992). The assumption of a theme position and floating phonological features in Pallarese gender exponence allows a simple treatment of 3rd person singular accusative masculine clitics and their use of *-o* for phonotactic reasons, since it has been shown in section 4.1.2 that right-edge epenthesis is not a synchronic phenomenon in the phonology of Catalan: the use of regular epenthesis is discarded in word-final position (the location for inflection) to avoid

mismatches between phonological and morphological structure. Moreover, the proposal also predicts that initial epenthesis (i.e., *[el] instead of [lo] in (5-b)) will be dispreferred because it implies the creation of more structure, that is, the insertion of new features and a skeletal position; both things are part of the input —[+labial] + V-slot— for the output [lo], and thus preferred over *[el].

The avoidance of initial epenthesis is not only true for the masculine singular but also for 3rd person accusative plural clitics (and definite articles), which make no gender distinctions; that is, both masculine and feminine clitics show the same phonological shape. However, as in (5), there are two phonetic variants *ls* and *les* that are phonologically conditioned, although the quality of the inserted vowel differs. This is illustrated in (7):

(7)    a.    No   **[ls]**             porto,   les         gerres        grogues
            NEG ACC.3PL.FEM bring.1S DEF.FEM.PL vase.FEM.PL yellow.FEM.PL
            'I don't bring them, the yellow vases'

        b.    No   **[ls]**             porto,   les         gots
            NEG ACC.3PL.MASC bring.1S DEF.MASC.PL glass.MASC.PL
            grocs
            yellow.MASC.PL
            'I don't bring them, the yellow glasses'

        c.    **[les]**           porto,   les         gerres        grogues
            ACC.3PL.FEM bring.1S DEF.FEM.PL vase.FEM.PL yellow.FEM.PL
            'I bring them, the yellow vases'

        d.    **[les]**           porto,   les         gots         grocs
             ACC.3PL.MASC bring.1S DEF.MASC.PL glass.MASC.PL yellow.MASC.PL
            'I bring them, the yellow glasses'

In (7-a) and (7-b), the clitic can be syllabified with the preceding vowel of its host, the adverb *no*. In absolute initial position before a vowel or a consonant —(7-c) and (7-d)—, by contrast, the vowel *-e*, which is the default unmarked vowel in Pallarese, is inserted. The difference in exponence with regard to the singular clitic in (5), which inserts *-o* in the same environment, is attributed to the lack of gender in the final morphosyntactic output prior to Vocabulary Insertion in the plural clitic, formally expressed by the postsyntactic Impoverishment rule depicted in (8) (the morph *-l* in the clitic is the realization of definiteness, i.e., [+def]):

(8)    $[\pm\text{fem}] \longrightarrow \emptyset \ /[+\text{def}]\_\_\_[+\text{pl}]$

After Impoverishment in the plural (a marked enviornment, see e.g. Bonet 1991, Noyer 1992 or Arregi & Nevins 2007), the clitic does not contain gender in its featural inventory, and thus no floating phonological features associated with gender

can be present in the input and used as a repair strategy. Even in the absence of gender features, though, the theme position is maintained (the underlying form for the plural clitic is /l–V+z/). Since the theme position with a V-slot is present in the input, the vowel -*e* is not considered fully epenthetic because only the features of the unmarked vowel of Pallarese have been inserted, not the structural position. This vowel is thus interpreted as thematic, although it has epenthesized its place features, and preferred over a full initial epenthetic vowel (i.e., *\*els*) that needs to create an X-slot.

In this dissertation, the dative is considered (following Martín 2012) a complex structure that consists of accusative + deixis, which implies that is also a marked structure like the plural (it is a superset of the accusative). Therefore, the Impoverishment rule has been extended to dative clitics and rephrased as in (9):

(9)     $[\pm\text{fem}] \longrightarrow \emptyset/[+\text{def}]\underline{\phantom{xx}} \left\{ \begin{array}{c} [+\text{pl}] \\ [+\textbf{dx}] \end{array} \right\}$

The presence of deixis triggers deletion of gender features in the dative. In the singular, *li* is found in all contexts for both feminine and masculine ([**li**] *toques la mà* 'You touch her/his hand'; *No* [**li**] *toques la mà* 'You don't touch her/his hand']. In the plural, the alternation *ls/les* is also phonologically conditioned (*No* [**ls**] *toques la mà* 'Don't touch their hand' vs. [**les**] *toques la mà* 'You touch their hand'), as for the accusative plural in (7). The process that inserts the place features of -*e* in the theme position of the accusative plural is also the same in the dative plural.

Pronominal clitics other than those referring to the 3rd person have been formally analyzed in chapter 6, where it has been shown that the theme position is also crucial in the final exponence they have. The interactions between epenthesis and the theme position are illustrated in (10) with the partitive clitic *ne*:

(10)     a.   [**n**]' agafa   quatre
             PART take.3S four
             'S/he takes four of that'

         b.   D' aguest  [**ne**] vull     tres
             of this.M.S PART want.1S three
             'I want three of this'

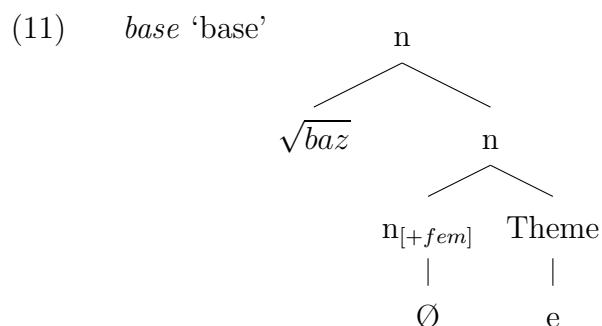The phonetic variants of the partitive are also phonologically conditioned, and the underlying form of the clitic, /n–V/, determines the location of the vowel that is needed for syllabification purposes in (10-b) (the partitive is syllabified with the vowel-initial verb in (10-a)). Gender is absent from the partitive morphosyntactic composition, and thus the input does not contain floating phonological features

either, as in the 3rd person accusative (plural) and dative (singular and plural), although in these pronouns this was due to Impoverishment. Again, the vocalic slot of the morphological theme position allows the insertion of place features and avoids initial epenthesis (*en*), which is more costly because it needs to create a skeletal position.

The same explanation of the interaction between epenthesis (of place features) and the theme position holds for the reflexive and 1st and 2nd person singular clitics, and also for the locative and neuter clitics (analyzed in section 6.3 with slight modifications). In addition, a tentative proposal for 1st and 2nd person plural clitics (which pose more problems for interpretation) has been given in 6.2.

All classes of nouns and adjectives have been analyzed in chapter 7. In regular feminine nouns like *cas*[a] 'house', -*a* always surfaces due to the ranking in (4), but in the masculine the same ranking favors outputs with unparsed features like *gos*[Ø] 'dog'. However, phonotactic requirements can make the floating phonological features of the input surface in the output, precisely in roots that end in a sibilant like *gos*, where the addition of the plural morph /z/ would yield an illicit sequence \**goss*. The floating features of the input are thus parsed in the plural *goss*[o]*s* 'dogs', showing once more that morphological structure constrains phonological epenthesis, since the same repair strategy used in 3rd person singular accusative masculine clitics (i.e., the use of an inflectional morph in stem-final position) is active in OCP-avoidance plurals like *goss*[o]*s*.

A consequence of the lexical entries posited for [±fem] in (1-b) is that vowels other than -*a* (fem.) or -*o* (masc.) cannot be considered gender exponents. The vowels that appear in non-regular elements must be specified in the theme position instead, as shown in (11):

(11)    *base* 'base'

$$
\begin{array}{c}
\text{n} \\
\diagup \diagdown \\
\sqrt{baz} \quad\quad \text{n} \\
\diagup \diagdown \\
\text{n}_{[+fem]} \quad \text{Theme} \\
| \quad\quad\quad | \\
\text{Ø} \quad\quad \text{e}
\end{array}
$$

In (11) the feminine word *base* does not bear the default feminine -*a* ending but -*e* instead. This vowel is not related to gender in any way, but it is nevertheless inflectional as it occupies the theme position. The specification of the theme vowel implies that the whole morphosyntactic structure in (11) is realized in one step by

/baze/, thus supporting multiple-terminal insertion (or phrasal spell-out) proposals, (e.g., Caha 2009, Svenonius 2012, Bermúdez-Otero 2013 or Merchant 2015). The storage of these complex structures (only in *non-regular* nominal elements) can dispense with the use of class markers in roots (Harris 1985 and subsequent work), which are problematic for the reasons given in sections 1.1 and 7.1.

By assuming a complex morphological structure with a general theme position in the nominal system, on the one hand, and an autosegmental spell-out process for gender, on the other hand, the analysis presented in this dissertation satisifes modularity, since there is no access to morphosyntactic information in the phonological module. Only the phonological features associated with gender and the underspecified vowel of the theme position are visible by the phonology, and all the constraints that play a role in the final exponence of clitics, nouns and adjectives refer exclusively to phonological objects such as association lines or phonological features. Furthermore, there is no need for gender allomorphy (cf. Bonet et al. 2007) if theme vowels are specified in the lexical entries of non-regular nominal elements, thus simplifying the computational system.

## 8.2  Microvariation: Central Catalan vs. Pallarese

It has been assumed throughout this dissertation that the morphologically salient theme position of the nominal system constrains phonological epenthesis in Pallarese Catalan. In case of phonological conflict, the grammar has two options to avoid the insertion of a new vowel:

1. If gender is a part of the morphosyntactic composition of a nominal element, the floating featural exponents associated with gender are available as a repair strategy and linked to the theme position *via* association lines. This is what happens, for example, in the masculine singular definite article alternation [l]/[lo], where *-o* surfaces only when forced by phonotactics (*que'* [l] *pa és bo* 'that the bread is good' vs. [lo] *pa és bo* 'the bread is good' *[l p]*a és bo*).

2. If the floating features associated with gender are not available (due to Impoverishment or because gender is not expressed in a syntactic category), the place features of the default vowel of Pallarese, *-e*, are epenthesized and linked to the placeless vocalic slot spelled out in the theme position (i.e., Th ↔ V), as in the alternation *ls/les* in the plural definite article (*que'* [ls] *pans són bons* 'that the bread loaves are good' vs. [les] *pans són bons* 'the bread loaves

are good' *[ls p]*ans són bons*). As argued in chapter 5, this is preferred over epenthesis of a full vowel with its own skeletal position.

The ideas mentioned above have been used for the analysis of the clitic system of Pallarese Catalan —the core of this thesis— but the proposal about the exponence of gender in clitics has been extended to other nominal categories in chapter 7. While the analysis of nouns and adjectives proposed for Pallarese can be applied to Central Catalan without major modifications (besides, obviously, the differences in vocalic quality due to vowel reduction) the realization of pronominal clitics (and definite articles) and their possible epenthetic loci pose some problems for interpretation.

It has been shown in previous chapters that Pallarese avoids initial (i.e., left-edge) epenthesis in clitics in all cases. In Central Catalan, by contrast, this kind of epenthesis is found in almost all pronominal proclitics in preconsonantal absolute position (or following a consonant-final word in preconsonantal position), as shown in table 8.1:[1]

| Label | Citation form | Contextual variants |
|---|---|---|
| 1st sg. Acc./Dat. | em | em, m', 'm, me |
| 2nd sg. Acc./Dat. | et | et, t', 't, te |
| 1st pl. Acc./Dat. | ens | ens, 'ns, nos |
| 2nd pl. Acc./Dat. | us | us, vos |
| 3rd Acc. sg. masc. | el | el, l', 'l, lo |
| 3rd Acc. sg. fem. | la | la, l' |
| 3rd Acc. pl. masc | els | els, 'ls |
| 3rd Acc. pl. fem. | les | les |
| 3rd sg./pl. reflex. | es | es, s', 's, se |
| 3rd Dat. sg. | li | li |
| 3rd Dat. pl. | els hi | els hi, 'ls hi |
| Neuter | ho | ho |
| Partitive | en | en, n', 'n, ne |
| Locative | hi | hi |

Table 8.1: *Central Catalan pronominal clitics*

---

[1]There is a lot of dialectal variation within Central Catalan. The clitic system depicted in table 8.1 belongs to a variety which is closer to Standard Catalan. For an intradialectal analysis of clitics in Central Catalan from a morphophonological perspective see Bonet & Lloret (2002).

Except for 3rd person accusative feminine (singular and plural), 2nd person plural, dative singular, locative and neuter clitics, all other pronouns can undergo initial epenthesis, which is phonetically [ə] in Central Catalan, e.g., *em* [əm]. The different epenthetic and non-epenthetic variants found for the 1st person singular clitic *em* are shown in (12):

(12)    a.    [**əm**]    compro  una    casa
                REFL.1S buy.1S  a.FEM.S house.FEM.S
                'I buy a house'

          b.    Estan  vigilant-    [**mə**]
                be.3PL guard.GER ACC.1S
                'They are watching me'

          c.    Porta'      [**m**]    a  casa!
                bring.IMPR.2S ACC.1S to house.FEM.S
                'Take me home!'

          d.    [**m**]'    agrada la      xocolata
                DAT.1S like.3S the.FEM.S chocolate.FEM.S
                'I like chocolate'

The 1st person singular clitic is assumed to be underlyingly /m/ in Central Catalan (Mascaró 1986, among others). This underlying form has also been posited for Pallarese Catalan in chapter 6 with a slight —but crucial— modification: a theme position with an underspecified vowel is attached to it, i.e., /m–V/. On the surface, the main difference between Pallarese and Central Catalan is the asymmetric use of epenthesis in the latter, which is found word-initially in (12-a) but word-finally in (12-b) instead. There is one single output in Pallarese Catalan in these same contexts, namely [me], as illustrated in (13):

(13)    a.    [**me**]    compro  una    casa
                REFL.1S buy.1S  a.FEM.S house.FEM.S
                'I buy a house'

          b.    Estan  vigilant-    [**me**]
                be.3PL guard.GER ACC.1S
                'They are watching me'

In both (13-a) and (13-b) *-e* is considered a theme vowel which has epenthesized place features, since the presence of a V-slot in the theme position licenses the insertion of default phonological features. It is precisely the presence of this position which bans the use of initial epenthesis, and therefore a form like *[em] —in which, besides

place features, an X-slot has also been epenthesized— is not possible in Pallarese.[2]

As for pre- or postvocalic clitics, as illustrated in (12-d) and (12-c), there is no need for a vowel to syllabify the phonological string, as the clitic can be attached to a preceding or following vowel and thus it surfaces as [m] in both dialects.

Given the differences between Central Catalan and Pallarese, one question arises: why do enclitics and proclitics pattern in different ways in Central Catalan? If the theme position, which is pervasive in the whole nominal system, is available in enclitic position, as in (12-b), why don't we find sequence like *[mə] *compro una casa* 'I buy a house'? That is, the nominal structure posited for Pallarese predicts that, when applied to Central Catalan, forms like [mə] should be favored in all circumstances (not only in enclitic position). Initial epenthesis should only be a last resort strategy: in the OT analysis of the 1st person singular clitic presented in chapter 6, any candidate with inital epenthesis is harmonically bound by a candidate with epenthesis of [–back] in the theme position because DEP-[–BACK] is a subset of DEP-V, a constraint that goes against the inserton of -*e*.

This question is particularly intriguing when confronted with the 3rd person singular masculine accusative clitic, due to empirical evidence that a morphological position of some kind must be assumed for it. In enclitic position, the inserted vowel is -[u], which is clearly an inflectional ending. It was shown in chapter 4.1.2 that word-final epenthesis is not a synchronic process of Catalan and is avoided because it interferes with the location for inflection. This is precisely what we see in (14) for the 3rd person singular masculine accusative clitic in Central Catalan:

(14)   a.   [əl]          compro
              ACC.3S.MASC buy.1S
              'I buy it'

       b.   Estan   vigilant-   [lu]
              be.3PL guard.GER ACC.3S.MASC
              'They are watching him'

       c.   Porta'        [l]          a  casa!
              bring.IMPR.2S ACC.3S.MASC to house.FEM.S
              'Take him home!'

       d.   [l]'          agafo
              ACC.3S.MASC take.1S
              'I take it'

---

[2]Recall that the default epenthetic vowel of Pallarese is -[e], not -[ə], and therefore the variant with initial epenthesis for the 1st person singular clitic should be *[em].

The use of this pronoun in proclitic and enclitic position is also asymmetric in Central Catalan.[3] In pre- or postvocalic positions, the clitic, which is underlyingly /l/, can be attached to a vocalic host, and thus there is a faithful input-output mapping —(14-c) and (14-d). In absolute initial position before a consonant, however, it surfaces with initial epenthesis, as in (14-a), whereas in postconsonantal enclitic position it surfaces as [lu], as in (14-b). Recall that 3rd person accusative clitics bear gender features (unlike 1st and 2nd person clitics) and that -[u] (-[o] in Pallarese) is the second most common inflectional ending for the masculine. It was proposed in chapter 5 that in Pallarese the features associated with masculine gender are usually not linked to the theme position due to the constraint ranking DEPLINK >> PARSE[+LABIAL] (the insertion of association lines for floating fatures is dispreferred), but sometimes syllabification problems force the linking of [+labial], which makes the vowel -*o* surface, as shown in (15):

(15)    a.    **[lo]**      compro
               ACC.3S.MASC buy.1S
               'I buy it'

        b.    Estan  vigilant-   **[lo]**
               be.3PL guard.GER ACC.3S.MASC
               'They are watching it'

In Pallarese, as shown in (15), the clitic always surfaces as [lo] when there is a phonotactic problem (*[l c]*ompro* – *vigila*[n l]). Therefore, the use of the features associated with [–fem] to repair syllabic structure is systematic. In Central Catalan, by contrast, the masculine exponent appears only in enclitic position, as in (14-d), whereas initial epenthesis is preferred in proclitics, as in (14-a). Again, the location of the epenthetic vowel differs depending on the clitic position (as in the 1st person singular clitic alternation [əm]/[mə]), although in the 3rd person singular accusative clitic this vowel is, word-finally, an exponent of gender. If, as is assumed for Pallarese, we consider that the clitic has a theme position also in Central Catalan, i.e., /l–V/, the linking of the features associated with [–fem] should be favored in all instances, thus generating outputs like *[lu] *compro* and not [əl] *compro*, as in (14-a).

    Interestingly, it is not only Central Catalan that patterns in this way. Cardinaletti & Repetti (2008) report similar behavior in Paduan subject clitics, exempli-

---

[3]The plural counterpart for the 3rd person singular masculine accusative clitic, which simply attaches the plural morph /z/ to /l/ (i.e., /l+z/), also shows the same asymmetry. In enclitic positions we find -[lus] when there is a syllabification problem, as in *vigilant-***[lus]** 'watching them', whereas in proclitic position its form is [əls], as in [əls] *compro* 'I buy them'. Therefore, whatever is assumed for the clitic /l/ is also valid for the plural /l+z/.

fied in (16):

(16)     /l-maɲa/ > [e̯l-máɲi]     'He eats'

         /maɲa-l/ > [máɲe-lo̯]     'Does he eat?'

Cardinaletti & Repetti (2008) argue that, in (16), the 3rd person singular clitic undergoes -*e* epenthesis (the usual epenthetic vowel of Paduan) in proclitic position, while the enclitic surfaces as [lo] because the phrase final position is morphologically salient in Paduan, and thus a morphologically neutral vowel is preferred:

> Why are two different epenthetic vowels used in proclitic and enclitic
> positions in (22a) and (22b)=(23c)? Since final position is reserved for
> (vocalic) inflectional morphemes, the epenthetic vowel in (23c) is in a
> morphologically salient position. And /e/ is a morphologically marked
> vowel in nominals that represents plural and feminine, two marked cate-
> gories in the (pro)nominal system. (23b) has a morphologically marked
> vowel (/e/) in a morphologically salient position (final position), an un-
> desirable structure given that [e] is epenthetic and therefore does not
> represent a morpheme. Instead, a *morphologically* neutral vowel is used
> in final position: /o/ (23c). (Cardinaletti & Repetti, 2008, p. 537)[4]

Cardinaletti and Repetti consider this vowel 'neutral' because it also appears phrase-finally in other pronouns where -*o* cannot be linked to masculine gender features, as in the 2nd person singular clitic: /t-maɲi/ > [te̯-máɲi] 'You eat' vs. /maɲi-t/ > [máɲi-to̯] 'Do you eat?'.[5] This vowel is not considered a morpheme but the default epenthetic *morphological* vowel of Paduan (as opposed to the default *phonological* one, -*e*), which receives "default interpretation for gender (masculine) and number (singular)" (p. 537). Forms like [máɲi-to̯] and [máɲe-lo̯] thus constitute cases of morphological epenthesis (see section 4.1.2).

Even if the motivations for the use of -*e* or -*o* in Paduan are different from those of Central Catalan (in the enclitic [lu], -[u] is considered the realization of mascu-line gender in Central Catalan, not a morphologically neutral vowel that receives default interpretation), what is undeniable is that there is a clear parallelism in the proclitic/enclitic alternation *el*/*lo* of the 3rd person clitic in both languages.[6] There

---

[4]In Cardinaletti & Repetti (2008), (22a) refers to the proclitic [e̯l-máɲi] and (22b)=(23c) to the enclitic [máɲe-lo̯], depicted in (16). (23b) refers to the non-existent form *[máɲe-le̯].

[5]Only the phrase final position is considered morphologically salient, and thus -*e* can appear at the right edge of the 2nd person singular clitic before the verb. The form *[et], parallel to [el] in the 3rd person, is not found because Paduan does not allow [t] in coda position.

[6]Moreover, Cardinaletti & Repetti (2008, p. 538) note shortly in footnote 19 that the analysis posited for Paduan subject clitics could be extended to Veronese object cltics: "In Veronese, the

is no doubt that morphological structure plays some role in determining the quality of the vowel, but the theme position advocated in this dissertation for the nominal system makes incorrect predictions for the surface forms of Central Catalan (and Paduan as well). As already noted, inputs like those of (17) should favor outputs with the features of -[u] associated with the V-slot of the theme position in all cases, but [lu] is only found in enclitic position:

(17)     /l–V#porto/     >     *[lu pɔ́rtu]  'I bring it'     (cf. [əl pɔ́rtə])

         /pɔrtaɾ#l–V/     >     [purtár lu]  'to bring it'

A possible solution to the enclitic/proclitic asymmetry of Central Catalan could be attributed to the constraint CONTIGUITY, as defined in (18) (McCarthy & Prince, 1995):

(18)     CONTIGUITY

         a.   I-CONTIGUITY ("No skipping")
              The portion of $S_1$ standing in correspondence forms a contiguous string.

         b.   O-CONTIGUITY ("No intrusion")
              The portion of $S_2$ standing in correspondence forms a contiguous string.

If we assumed an input /l#pɔrto/ with no theme position (and, thus, no underlying V-slot), CONTIGUITY, which prevents internal epenthesis (and also deletion), would rule out a candidate like *[lu̱ pɔ́rtə], as it would break up the phonological string. This is shown in (19):

(19)     [əl] *porto* 'I bring it'

| /l # p ɔ r t o/ | CONT | $\sigma$STR | DEP-V |
|---|---|---|---|
| a.   l  u    p  ɔ́  r  t  u | 1W | | 1 |
| b.   l      p  ɔ́  r  t  u | | 1W | L |
| c. ☞  ə  l    p  ɔ́  r  t  u | | | 1 |

Without a general theme position in the nominal system, it does not make sense to assume that the phonological features associated with gender are floating, since

---

3sg masculine accusative pronoun /l/ surfaces as [el] in proclitic position (*El so* 'I know it (lit. it [I] know)') and as [lo] in enclitic position (*magnar-lo* 'to eat it')." This is exactly what happens in Central Catalan, shown in (14-a) and (14-b).

there is no morphological position with an underspecified vowel to which those features can be attached. Thus, these features are not part of the input in (19), where candidate c., which inserts a vowel but nevertheless respects CONTIGUITY, is the winning candidate because candidate a. violates CONTIGUTIY and candidate b. is ruled out due to syllabification issues.[7] The situation changes in enclitic position, though, where *portar-l*[u] 'to take it' would have an input /pɔrtaɾ#l/ and therefore the insertion of -[u] in word-final position would respect CONTIGUITY.

Although an analysis like the one just sketched would work for Central Catalan, the different vowel quality of the epenthesized segment in enclitic and proclitic position is still difficult to account for if no theme position is assumed. That is, -[u] is an inflectional marker, but no general mechanism can explain why -[u] is inserted only word-finally. This is an undesirable consequence of proposals that posit a simpler morphological structure for the clitic (and the nominal system).

If, by contrast, as shown in (20), we assume a theme position, the same constraint ranking (with the addition of DEPLINK and PARSE[+LABIAL]) incorrectly predicts the output [lu]:[8]

(20)

| [+labial]<br><br>/l  V  #  p  ɔ  r  t  o/ | CONT | σSTR | DEP-V | DEP-LK | PARSE[+LAB] |
|---|---|---|---|---|---|
| [+labial]<br>\|<br>a. ☞  l  u    p  ɔ́  r  t  u | | | | 1 | |
| [+labial]<br><br>b.   l      p  ɔ́  r  t  u | 1W | | | L | 1 |
| [-back] [+labial]<br>\|<br>c. ☹  ə  l      p  ɔ́  r  t  u | | | 1W | 1 | 1 |

---

[7]Bonet & Lloret (2005) consider that CONTIGUITY only applies morpheme-internally ('IO-CONTIGUITY bans morpheme internal deletion or epenthesis', p.56), and thus candidate a. would not violate CONTIGUTIY in (19). It is assumed here instead that the constraint also bans intermorphemic epenthesis.

[8]The vowel -[u] associated with [–fem] bears the features [+labial,+high], but I simplify the tableau in (20) by just including [+labial] in the input.

As there are floating features in the input, the ranking in (20) favors *[lu] over [əl] because there is already a vocalic slot for the association of [+labial], and therefore candidate a. does not violate DEP-V. The constraint CONTIGUITY could rule out a form *[lu], but it is the theme position itself which makes candidate a. satisfy CONTIGUITY. The V-slot of the theme position is part of the morpheme /l–V/, and the association of -[u] to V maintains the clitic and the verb contiguous in both the input and output. If we assume a morphological theme position, there is no possible way in which a vowel that is located therein can be considered to be breaking up the phonological string.

Bonet & Lloret (2005) make use of alignment constraints to account for these differences, namely ALIGN(CL/V) ("Align the right edge of a pronominal clitic with the left edge of V").[9] Leaving aside, for the sake of the argument, that this constraint would contradict the *desideratum* of modularity advocated in this thesis, one could invoke alignment constraints that made specific reference to clitic and verb edges. Note, however, that this solution faces the same poblems shown above for input-output contiguity: the theme position is a part of the clitic, and therefore the vowel -[u] is aligned with the verb in the output *[lu] *porto*, thus satisfying ALIGN(CL/V). The only solution that could work with a theme position in the input would be to posit a constraint that asks for the definiteness morph to be adjacent to the verb (*[**lu**] *porto* would then violate both ALIGN(CL/V) and CONTIGUITY), but this clearly violates all principles of independence between morphosyntax and phonology and must therefore be discarded.

The interactions between syllabification and morphological structure seem more intricate in the clitic system of Central Catalan as the proclitic-enclitic asymmetries are difficult to account for by referring only to the theme position. A thorough analysis of this dialect that allows us to understand how to accommodate it to the theoretical proposal defended here falls beyond the scope of this dissertation and will thus be left for further research.

---

[9]Bonet & Lloret (2005) distinguish between ALIGN(CL/V) and CONTIGUITY due to the complex epenthetic patterns of Central Catalan clitics that complicate their morphophonological analyisis. For the present purposes, though, these differences are obviated.

# PARTICIPANTS

## JN

| | |
|---|---|
| Age | 79 |
| Gender | Male |
| Location | Son (Vall d'Àneu) |
| Mother | Son (Vall d'Àneu) |
| Father | Son (Vall d'Àneu) |
| Education | Primary school |
| Occupation | Shepherd (pensioner at the time of the interview) |
| Date of interview | May 2014 |
| Duration | 1:29 h |
| Observations | – |

## PE

| | |
|---|---|
| Age | 85 |
| Gender | Male |
| Location | València d'Àneu (Vall d'Àneu) |
| Mother | València d'Àneu (Vall d'Àneu) |
| Father | València d'Àneu (Vall d'Àneu) |
| Education | Primary school |
| Occupation | Shepherd (pensioner at the time of the interview) |
| Date of interview | May 2014 |

Duration             00:39 h

Observations         His wife (77, born in Isil, also in Vall d'Àneu) was present
                     during the interview and participated in the translation task
                     actively. They spent some periods of time with their
                     children in Barcelona.

# PP

Age                  80

Gender               Male

Location             Alós d'Isil (Vall d'Àneu)

Mother               Alós d'Isil (Vall d'Àneu)

Father               Alós d'Isil (Vall d'Àneu)

Education            Primary school

Occupation           Shepherd (pensioner at the time of the interview)

Date of interview    May 2014

Duration             00:47 h

Observations         –

# II

Age                  82

Gender               Male

Location             Àreu (Vall Ferrera)

Mother               Àreu (Vall Ferrera)

Father               Àreu (Vall Ferrera)

Education            Primary school

Occupation           Shepherd (pensioner at the time of the interview)

Date of interview    May 2014

Duration             00:52 h

Observations         –

# RA

| | |
|---|---|
| Age | 87 |
| Gender | Female |
| Location | Àreu (Vall Ferrera) |
| Mother | Àreu (Vall Ferrera) |
| Father | Àreu (Vall Ferrera) |
| Education | Primary school |
| Occupation | Farmer and housewife (pensioner at the time of the interview) |
| Date of interview | May 2014 |
| Duration | 00:47 h |
| Observations | She had spent a short part of her childhood in France. |

# RO

| | |
|---|---|
| Age | 87 |
| Gender | Female |
| Location | Àreu (Vall Ferrera) |
| Mother | Àreu (Vall Ferrera) |
| Father | Àreu (Vall Ferrera) |
| Education | Primary school |
| Occupation | Farmer and housewife (pensioner at the time of the interview) |
| Date of interview | May 2014 |
| Duration | 00:34 h |
| Observations | – |

# JE

| | |
|---|---|
| Age | 67 |
| Gender | Male |
| Location | Estaís (Vall d'Àneu) |
| Mother | Estaís (Vall d'Àneu) |
| Father | Estaís (Vall d'Àneu) |
| Education | Primary school |
| Occupation | Farmer and worker in a ski station |

| | |
|---|---|
| Date of interview | May 2014 |
| Duration | 00:35 h |
| Observations | – |

# CO

| | |
|---|---|
| Age | 69 |
| Gender | Female |
| Location | Farrera (Vall Ferrera) |
| Mother | Farrera (Vall Ferrera) |
| Father | Farrera (Vall Ferrera) |
| Education | Secondary school |
| Occupation | Hotel manager |
| Date of interview | May 2014 |
| Duration | 00:54 h |
| Observations | Originally from Farrera, but she had been living in Escaló since she was a teenager. She spoke French. A friend from hers (68, Alós d'Isil) was also present during the interview and participated in the translation task actively . |

# JD

| | |
|---|---|
| Age | 77 |
| Gender | Male |
| Location | Ainet de Besan (Vall Ferrera) |
| Mother | Ainet de Besan (Vall Ferrera) |
| Father | Ainet de Besan (Vall Ferrera) |
| Education | Primary school |
| Occupation | Farmer and worker in an electrical company (pensioner at the time of the interview) |
| Date of interview | May 2014 |
| Duration | 00:49 h |
| Observations | – |

# PA

| | |
|---|---|
| Age | 69 |
| Gender | Female |
| Location | Araós (Vall Ferrera) |
| Mother | Araós (Vall Ferrera) |
| Father | Araós (Vall Ferrera) |
| Education | Primary school |
| Occupation | Farmer and housewife |
| Date of interview | May 2014 |
| Duration | 00:38 h |
| Observations | Her husband (73, also from Araós) participated in the translation task during the interview. |

# CA

| | |
|---|---|
| Age | 88 |
| Gender | Female |
| Location | Araós (Vall Ferrera) |
| Mother | Araós (Vall Ferrera) |
| Father | Araós (Vall Ferrera) |
| Education | Primary school |
| Occupation | Farmer and housewife |
| Date of interview | May 2014 |
| Duration | 00:42 h |
| Observations | Originally from Araós, but she was living in Alins since 2012. |

# QUESTIONNAIRE

## Clitics in isolation

### 1st person singular

- No me gusta el chocolate

- Me parece correcto

- No me parece correcto

- Esto me parece bien

- Quiere llevarme a casa

- Están buscándome

- Tráeme la cena

- Traedme la cena

- Añádeme al grupo

- Cóseme el botón

### 2nd person singular

- Te cogen el teléfono

- Te sigue la policía

- Que te sigue la policía

- Esto, te quiere decir

- Quieren llamarte

- Estan llamándote

- Conéctate a esa línea

- Añádete al grupo

- Cósete el botón

## 1st person plural

- No nos interesa la política

- No nos cae bien

- Nos necesita

- Nos llama para llevarnos al cine

- Están vigilándonos

- Tráenos el pastel

- Traednos el pastel

- Comprémonos el pastel

- Añádenos a la lista

- Cósenos el botón

## 2nd person plural

- No os alejéis de aquí

- No os quedéis aquí

- Os necesita

- Quiere llevaros al cine

- Están vigilándoos

- Agarraos fuerte

### 3rd person singular dative

- Le coge la mano

- Le compras un regalo

- Cogerle la mano

- Dale un libro

- Démosle un libro

- Exígele eso

- Cósele el botón

### 3rd person plural dative

- Les cojo la mano

- Les traigo un regalo

- Quiere llevarles un regalo

- Cántales una canción

- Cantémosles una canción

- Exígeles que vengan

- Cóseles el botón

### 3rd person (singular and plural) masculine accusative

- La cama, la ensucian ellos

- La cama, la dejaré limpia

- La cama plegable, que la coja él

- La cama, que la dejen limpia

- Las camas, las ensucian ellos

- Las camas, las dejaré limpias

- Las camas plegables, que las cojan ellos

- Las camas, que las dejen limpias ellos

- No quería hacerla, esa cama

- No quería hacerlas, esas camas

- Haciéndola, la cama, antes que él, lo felictaron

- Haciéndolas, las camas, antes que él, lo felicitaron

- Hazla, esa cama

- Hazlas, esas camas

- Hacedla, esa cama

- Hacedlas, esas camas

## 3rd person (singular and plural) feminine accusative

- La tienda, la ensucian ellos

- La tienda, la dejaré limpia

- Las tiendas, las ensucian ellos

- Las tiendas, las dejaré limpias

- No podían encontrarla, la tienda

- No podían encontrarlas, las tiendas

- Encontrándola, le hicieron un favor

- Encontrándolas, le hicieron un favor

- Encuéntrala, esa libreta

- Encuéntralas, esas libretas

## 3rd person reflexive (singular and plural)

- Se interesa por la literatura

- Se interesan por la literatura

- Se sale de su sitio

- Se salen de su sitio

- Se cansa muy deprisa

- Se cansan muy deprisa

- Quería verse en casa

- Querían verse en casa

- Quería casarse en mayo

- Querían casarse en mayo

- Queriéndose casar en mayo, tuvo que esperarse a junio

- Queriéndose casar en mayo, tuvieron que esperarse a junio

- Hacerse daño

## Neuter

- Esto, lo coge para su casa

- Esto, lo compra en el súper

- Esto, quería comprarlo ahí

- Esto, comprándolo ahí, ahorrarás dinero

- cómpralo en esa tienda

- compradlo en esa tienda

- Esto, pártelo así

- Esto, cóselo así

## Locative

- Ahora voy

- A aquel armario, no llego

- Al súper, no quería ir

- Añadiéndole sal, la sopa mejoró

- Ponle sal, a la sopa

- Al vestido, cósele el botón

## Partitive

- Quiero cuatro

- El chico coge cuatro

- De las tres camisas, quiere comprar solo una

- Comprando cuatro te regalan uno

- De los cuatro que hay aquí, compra solo uno

- De los cuatro que hay aquí, comprad solo uno

- De los cuatro que hay aquí, compremos solo uno

- De los dos panes, parte solo uno

- De los dos botones, cose solo uno

# Combinations

## 1st person singular + 3rd person (singular and plural) masculine accusative

- La lechuga, me la coge cada día del huerto

- Las lechugas, me las coge cada día del huerto

- Me la trae cada día, la lechuga

- Me las trae cada día, las lechugas

- La lechuga, no quería regalármela

- Las lechugas, no quería regalármelas

- Llevándomela, la lechuga, ahorré dinero

- Llevándomelas, las lechugas, ahorré dinero

- Esa lechuga de ahí, tráemela

- Esas lechugas de ahí, tráemelas

## 1st person singular + 3rd person (singular and plural) feminine accusative

- La galleta me la coge del bolso

- Las galletas me las coge del bolso

- La galleta me la trae para luego

- Las galletas me las trae para luego

- Esa galleta, querían cogérmela

- Esas galletas, querían cogérmelas

- Dejándomela, esa galleta, me alegraron la tarde

- Dejándomelas, esas galletas, me alegraron la tarde

- Esa galleta, tráemela

- Esas galletas, tráemelas

## 2nd person singular + 3rd person (singular and plural) masculine accusative

- Te la coge cada día del huerto, la lechuga

- Te las coge cada día del huerto, las lechugas

- Te la trae cada día del huerto, la lechuga

- Te las trae cada día del huerto, las lechugas

- Esa lechuga no puedes dejártela

- Esas lechugas no puedes dejártelas

- Comiéndotela toda, esta lechuga, tendrás dolor de barriga

- Comiéndotelas todas, estas lechugas, tendrás dolor de barriga

- Esa lechuga, cómetela

- Esas lechugas, cómetelas

## 2nd person singular + 3rd person (singular and plural) feminine accusative

- El tomate te lo trae tu padre

- Los tomates te los trae tu padre

- El tomate te lo aliña tu padre

- Los tomates te los aliña tu padre

- Ese tomate, tienes que comértelo

- Esos tomates, tienes que comértelos

- El tomate, tienes que acabar comiéndotelo

- Los tomates, tienes que acabar comiéndotelos

- Ese tomate, cómetelo

- Esos tomates, cómetelos

## reflexive + 3rd person (singular and plural) masculine accusative

- La merluza, se la calienta un poco en la sartén

- Las merluzas, se las calientan un poco en la sartén

- La merluza, se la hace a la plancha

- Las merluzas, se las hace a la plancha

- La merluza, quería añadírsela a su propio menú

- Las merluzas, quería añadírselas a su propio menú

- La merluza, le gustó más hirviéndosela

- Las merluzas, le gustaron más hirviéndoselas

## reflexive + 3rd person (singular and plural) feminine accusative

- El tomate, se lo lleva cada día de su huerto

- Los tomates, se los lleva cada día de su huerto

- El tomate, se lo come en la ensalada

- Los tomates, se los come en la ensalada

- EL tomate, le gusta comérselo con aceite y sal

- Los tomates, le gusta comérselos con aceite y sal

- El tomate, acabó comiéndoselo después

- Los tomates, acabó comiéndoselos después

## 1st person plural + 3rd person (singular and plural) masculine accusative

- La merluza, nos la acerca todos los martes

- Las merluzas, nos las acercan todos los martes

- La merluza, no nos la trae al restaurante

- Las merluzas, no nos las trae al restaurante

- La merluza, tenéis que acercárnosla

- Las merluzas, tenéis que acercárnoslas

- Cocinándonosla, han gastado toda la merluza

- Cocinándonoslas, han gastado todas las merluzas

- La merluza, acercádnosla

- Las merluzas, acercádnoslas

- La merluza, comprémonosla en el mercado

- Las merluzas, comprémonoslas en el mercado

## 1st person plural + 3rd person (singular and plural) feminine accusative

- El tomate nos los aliñamos poco

- Los tomates nos los aliñamos poco

- El tomate no nos lo quiere comprar nadie

- Los tomates no nos los quiere comprar nadie

- Para aliñárnoslo bien, el tomate, hace falta más aceite

- Para aliñárnoslos bien, los tomates, hace falta más aceite

- Acabaremos con todo el aceite, aliñándonoslo tanto, el tomate

- Acabaremos con todo el aceite, aliñándonoslos tanto, los tomates

- Ese tomate, aliñádnoslo

- Esos tomates, aliñádnoslos

## 2n person plural + 3rd person (singular and plural) masculine accusative

- La merluza, os la acerca al restaurante

- Las merluzas, os las acerca al restaurante

- La merluza, os la trae al restaurante

- Las merluzas, os las traen al restaurante

- La merluza, tenéis que comprarosla

- Las merluzas, tenéis que compraroslas

- Acabaremos comprándoosla, la merluza

- Acabaremos comprándooslas, las merluzas

- Esa merluza, coméosla

- Esas merluzas, coméoslas

## 2nd person plural + 3rd person (singular and plural) feminine accusative

- El tomate no os lo aliñáis

- Los tomates no os los aliñáis

- El tomate no os lo compréis aquí

- Los tomates no os los compréis aquí

- Compró el tomate aquí para regalároslo

- Compró los tomates aquí para regalároslos

- Me hacéis un favor comiéndooslo aquí, el tomate

- Me hacéis un favor comiéndooslos aquí, los tomates

- Compráoslo aquí, el tomate

- Compráoslos aquí, los tomates

## 3rd person singular dative + 3rd person (singular and plural) masculine accusative

- La merluza, se la llevo los lunes, a él

- La merluza, se la entrego los lunes, a él

- Las merluzas, se las llevo los lunes, a él

- La merluzas, se las entrego los lunes, a él

- La merluzas, dice que se las lleva los lunes, a él

- La merluzas, dice que se las entrega los lunes, a él

- La merluza, tiene que llevársela los lunes, a él

- Las merluzas, tiene que llevárselas los lunes, a él

- Dándosela a ella, la merluza, lo solucionaremos

- Dándoselas a ella, las merluzas, lo solucionaremos

- La merluza, démosela a ella

- Las merluzas, démoselas a ellas

## 3rd person singular dative + 3rd person (singular and plural) feminine accusative

- El tomate, se lo llevo los martes, a él

- El tomate, se lo entrego los martes a ella

- Los tomates, se los llevo los martes, a ella

- Los tomates, se los entrego los martes, a él

- Los tomates, dice que se los lleva los martes, a él

- Los tomates, dice que se los entrega los martes, a ella

- El tomate, tiene que llevárselo los lunes, a él

- Los tomates, tiene que llevárselo los lunes, a él

- Dándoselo a él, el tomate, lo solucionaremos

- Dándoselos a él, los tomates, lo solucionaremos

- El tomate, démoselo a ella

- Los tomates, démoselos a él

## 3rd person plural dative + 3rd person (singular and plural) masculine accusative

- La merluza, se la llevo los lunes, a ellos

- La merluza, se la entrego los lunes, a ellas

- Las merluzas, se las llevo los lunes, a ellas

- Las merluzas, se las entrego los lunes, a ellos

- Las merluzas, dice que se las lleva los lunes, a ellas

- Las merluzas, dice que se las entrega los lunes, a ellos

- La merluza, tiene que llevársela los lunes, a ellos

- Las merluzas, tiene que llevárselas los lunes, a ellas

- Dándosela a ellas, la merluza, lo solucionaremos

- Dándoselas a ellas, las merluzas, lo solucionaremos

- La merluza, démosela a ellas

- Las merluzas, démoselas a ellas

## 3rd person plural dative + 3rd person (singular and plural) feminine accusative

- El tomate, se lo llevo los martes, a ellos

- El tomate, se lo entrego los martes, a ellos

- Los tomates, se los llevo los martes, a ellas

- Los tomates, se los entrego los martes, a ellos

- Los tomates, dice que se los lleva los martes, a ellos

- Los tomates, dice que se los entrega los martes, a ellos

- El tomate, tiene que llevárselo los lunes, a ellas

- Los tomates, tiene que llevárselo los lunes, a ellas

- Dándoselo a ellos, el tomate, lo solucionaremos

- Dándoselo a ellos, los tomates, lo solucionaremos

- El tomate, démoselo a ella

- Los tomates, démoselos a él

### 1st person singular + partitive

- Me voy de aquí

- Zapatos, me compro un par

- De esas, me cojo una

- Irme de aquí

- Zapatos, quiero comprarme un par

- Tostadas, estoy tostándome un par

- Tostadas, tóstame un par

- Tostadas, sírveme un par

### 2nd person singular + partitive

- Te vas de aquí

- Zapatos, te compro cuatro

- De esas, te cojo una

- Irte de aquí

- Zapatos, quiero comprarte un par

- Tostadas, estoy tostándote un par

- Tostádas, cómete dos

- Tostadas, sírvete un par

### reflexive + partitive

- Se va de aquí

- Zapatos, se compra un par

- De esas, se coge una

- Irse de aquí

- Zapatos, quiere comprarse un par

- Tostándose cuatro, (de) tostadas, tardará mucho

## 1st person plural + partitive

- Nos vamos de aquí

- Zapatos, nos compramos un par

- De esas, nos cogemos una

- Irnos de aquí

- Zapatos, tenemos que comprarnos un par

- Tostadas, estamos tostándonos un par

- Vámonos

- Tostadas, sírvenos un par

## 2nd person plural + partitive

- Os vais de aquí

- De esas, os cogéis cuatro

- Zapatos, os compráis un par

- Para iros de aquí, tenéis que salir por allí

- Zapatos, tenéis que compraros cuatro

- Tostadas, estáis tostándoos cuatro

- Idos de aquí

## 3rd person singular dative + partitive

- A María le doy tres rosas. A ella, le doy dos

- A María le envío tres rosas. A ella, le envío dos

- A María quiero darle cuatro rosas. A ella, quiero darle dos

- A María quería darle cuatro rosas pero acabó dándole dos

- A María dale tres, (de) rosas

- A María ofrécele dos, (de) rosas

### 3rd person plural dative + partitive

- A María y a Julia les doy tres rosas. A sus hermanas, les doy cuatro

- A María y a Julia les envío tres rosas. A sus hermanas, les envío cuatro

- A María y a Julia quiero darles cuatro rosas. A sus hermanas, quiero darles seis

- A María y a Julia quería darles cuatro rosas pero acabó dándoles dos

- A tus hermanas, dales tres, (de) rosas

- A tus hermanas, ofréceles tres, (de) rosas

### 1st person singular + neuter

- Esto, me lo compra siempre en esa tienda

- Esto, me lo coge del huerto especialmente para mí

- No hace falta comprármelo, eso

- Esto, vinieron explicándomelo por el camino

- Eso de ahí, cómpramelo ahora

### 2nd person singular + neuter

- Eso te lo dice él pero no es verdad

- Esto te lo esconde para que no lo veas

- Esto, comprártelo aquí me parece una tontería

- Esto, vinieron explicándotelo por el camino

- Eso, cómpratelo ahora

### reflexive + neuter

- Eso, se lo coge para ella (misma)

- Esto, se lo compra para él (mismo)

- Esto, quería comprárselo para ella (misma)

- Esto, vinieron recordándoselo a ellos mismos

### 1st person plural + neuter

- Esto, nos lo aparta ella para la boda

- Esto, nos lo dice ahora pero no es verdad

- Comprárnoslo aquí, esto, me parece una tontería

- Esto, vinieron cantándonoslo

- Esto, comprémonoslo ya

### 2nd person plural + neuter

- Esto, os lo aparta ella para la boda

- Esto, os lo dice ahora pero no es verdad

- Esto, comprároslo hoy no me parece bien

- Esto, vinieron cantándooslo por el camino

- Esto, lleváoslo ya de aquí

### 3rd person singular dative + neuter

- Esto, se lo coge a su hermana

- Esto, se lo da a su hermana

- Te dije que vinieras para demostrárselo

- Demostrándoselo, se lo va a creer

- Esto, demuéstraselo de una vez

### 3rd person plural dative + neuter

- Esto, a tus hijos, dáselo

- Esto, a tus hijos, se lo doy

### 1st person singular + locative

- A Girona, nunca me acompaña

- A Girona, a ver cuándo me lleváis

- A Girona, no quería llevarme

- A Girona, llevadme un día de estos

- A Girona, acabó llevándome cunado él quiso

### 2nd person singular + locative

- A las reuniones, nunca te acompaña

- A las reuniones, nunca te lleva

- Quieren llevarte un día de estos, a Girona

- Acercándote en coche, a Girona, te harán un favor

- A Girona, acércate esta semana

### reflexive + locative

- Allí no se esconde nadie

- Aquí, hoy no se quiere quedar

- Todo es ponerse

### 1st person plural + locative

- En el castillo, no nos esconderemos

- Aquí, hoy no nos queremos quedar

- Para acercarnos tuvimos que ir en silencio

- Acercándonos, no conseguiremos nada

- Escondámonos, en el castillo

## 2nd person plural + locative

- Aquí, hoy no os quedaréis

- En mi casa, yo no os espero nunca

- Se acercó al aeropuerto para esperaros cuando salierais

- En la entrada, estuvieron tres horas esperándoos

## 3rd person (singular and plural) feminine accusative + locative

- El tomate, en la nevera, no lo pondré

- El tomate, en la nevera, ponlo

- Los tomates, en la nevera, no los pondré

- Los tomates, en la nevera, ponlos

## 3rd person (singular and plural) masculine accusative + locative

- La merluza, en la nevera, no la pondré

- La merluza, en la nevera, ponla

- Las merluzas, en la nevera, no las pondré

- Las merluzas, en la nevera, ponlas

## partitive + locative

- ¿Cuántos quedan en el bote? Quedan tres

- ¿Cuántos hayen el bote? Hay tres

- En el estofado, poner un poco más, (de) sal, no estaría mal

- En el estofado, acabó poniendo demasiada, de sal

- Sal, pon un poco más, en el estofado

### reflexive + 1st person singular

- Se me cierra el estómago al ver eso

- Se me encoge si lo pongo en la lavadora

- Para rompérseme así no se qué tuve que hacer

- Rompiéndoseme así no podré terminarlo nunca

### reflexive + 2nd person singular

- Se te cierra el estómago al ver eso

- Se te encoge si lo pones en la lavadora

- Para rompérsete así, tuvo que caer desde muy arriba

- Rompiéndosete así no podrás terminarlo nunca

### reflexive + 1st person plural

- Se nos cierra el estómago al ver eso

- Se nos encogen si los pones en la lavadora

- Al inundársenos la barca tuvimos que abandonar

- Rompiéndosenos así no podremos terminarlo nunca

### reflexive + 2nd person plural

- Se os cierra el estómago al ver eso

- Se os encogen si los ponéis en la lavadora

- Al inundárseos la barca tuvisteis que abandonar

- Rompiéndoseos así no podréis terminarlo nunca

### reflexive + 3rd person singular dative

- Se le cierra el estómago al ver eso

- Se le llena la boca con ese tema

- Al llevársele a su hijo, se quedó desconsolado

- Rompiéndosele así no conseguirá terminarlo nunca

### reflexive + 3rd person plural dative

- Se les cierra el estómago al ver eso

- Se les llena la boca con ese tema

- Al llevárseles a su hijo, se quedaron desconsolados

- Rompiéndoseles así no conseguirán terminarlo nunca

# BIBLIOGRAPHY

Acquaviva, P. (2009). Roots and lexicality in Distributed Morphology. In Galani, A., Redinger, D., & Yeo, N. (Eds.), *Proceedings of the "Fifth York-Essex Morphology Meeting" (YEMM)*. York Papers in Linguistics.

Akinlabi, A. (1996). Featural affixation. *Journal of Linguistics*, *32*(02), 239–289.

Alcover, A. M. & Moll, F. d. B. (1933/2011). *La flexió verbal en els dialectes catalans*. Online database.

Alexiadou, A. (2001). *Functional structure in nominals: nominalization and ergativity*. Amsterdam: John Benjamins.

Alexiadou, A. (2004). Inflection class, gender and DP internal structure. In G. Müller, L. Genkel, & G. Zifonoun (Eds.), *Explorations in nominal inflection* (pp. 21–49). Berlin: Walter de Gruyter.

Anderson, S. R. (1992). *A-morphous morphology*. Cambridge: Cambridge University Press.

Anderson, S. R. (2015). Words and paradigms: Peter H. matthews and the development of morphological theory. *Transactions of the Philological Society*.

Aronoff, M. (1994). *Morphology by itself: Stems and inflectional classes*, volume 22. Cambridge, MA: MIT press.

Arregi, K. & Nevins, A. (2007). Obliteration vs. Impoverishment in the Basque g-/z-constraint. *University of Pennsylvania Working Papers in Linguistics*, *13*(1), 2.

Arsenijević, B. (2009). Gender as a grammaticalized classifier system: the case of the Serbo-Croatian neuter. Available at http://ling.auf.net/lingbuzz/002848.

Badecker, W., Miozzo, M., & Zanuttini, R. (1995). The two-stage model of lexical retrieval: Evidence from a case of anomia with selective preservation of grammatical gender. *Cognition, 57*(2), 193–216.

Benveniste, E. (1966). *Problèmes de linguistique générale, tome 1.* Paris: Gallimard.

Bermúdez-Otero, R. (1999). *Constraint interaction in language change: quantity in English and Germanic.* PhD thesis, University of Manchester.

Bermúdez-Otero, R. (2006a). *Against nominal class features in Spanish.* Handout, *Workshop on Theoretical Morphology 2*, Universitat Leipzig, 17 June 2006.

Bermúdez-Otero, R. (2006b). Morphological structure and phonological domains in spanish denominal derivation. In F. M. Gil & S. Colina (Eds.), *Optimality-theoretic studies in Spanish phonology* (pp. 278–311). Amsterdam: John Benjamins.

Bermúdez-Otero, R. (2007). Spanish pseudoplurals: phonological cues in the acquisition of a syntax-morphology mismatch ricardo bermúdez-otero university of manchester. In M. Baerman, G. Corbett, D. Brown, & A. Hippisley (Eds.), *Deponency and morphological mismatches (Proceedings of the British Academy)* (pp. 231–269). Oxford: Oxford University Press.

Bermúdez-Otero, R. (2012). The architecture of grammar and the division of labour in exponence. In J. Trommer (Ed.), *The morphology and phonology of exponence* (pp. 8–83). Oxford: Oxford University Press.

Bermúdez-Otero, R. (2013). The Spanish lexicon stores stems with theme vowels, not roots with inflectional class features. *Probus, 25*(1), 3–103.

Bermúdez-Otero, R. (2016). We do not need structuralist morphemes, but we do need constituent structure. In H. Harley & D. Siddiqi (Eds.), *Morphological metatheory* (pp. 387–429). Amsterdam: John Benjamins.

Bobaljik, J. D. (2012). Distributed Morphology. *Ms. University of Connecticut.*

Boersma, P. (1997). How we learn variation, optionality, and probability. In *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam*, volume 21, (pp. 43–58).

Bonet, E. (1991). *Morphology after syntax: Pronominal clitics in Romance.* PhD thesis, Massachusetts Institute of Technology.

Bonet, E. (1995a). Feature structure of Romance clitics. *Natural language & linguistic theory*, *13*(4), 607–647.

Bonet, E. (1995b). The where and how of clitic order. *Révue Quebecoise de Linguistique-Université du Quebec a Montréal*, *24*, 61–82.

Bonet, E. (2002). Cliticització. In J. S. et al. (Ed.), *Gramàtica del català contemporani*, volume 1 (pp. 933–989). Barcelona: Empúries.

Bonet, E. (2006). Gender allomorphy and epenthesis in Spanish. In F. M. Gil & S. Colina (Eds.), *Optimality-theoretic studies in Spanish phonology* (pp. 278–311). Amsterdam: John Benjamins.

Bonet, E. (2008a). Item-and-arrangement or item-and-process? *Cuadernos de Linguística XV*, 1–12.

Bonet, E. (2008b). The Person-Case constraint and repair strategies. In R. D'Alessandro, S. Fischer, & G. H. Hrafnbjargarson (Eds.), *Agreement restrictions* (pp. 103–128).

Bonet, E. (2013). Variation in the exponence of determiners and other pronominal elements. *Presented at the Workshop on Syntactic Variation in Catalan and Spanish, Universitat Autòonoma de Barcelona, 26-28 June.*

Bonet, E. & Harbour, D. (2012). Contextual allomorphy. In J. Trommer (Ed.), *The morphology and phonology of exponence* (pp. 195–235). Oxford: Oxford University Press.

Bonet, E. & Lloret, M.-R. (1998). *Fonologia catalana.* Barcelona: Ariel.

Bonet, E. & Lloret, M.-R. (2002). OCP effects in Catalan Cliticization. *Catalan Journal of Linguistics*, *1*, 19–39.

Bonet, E. & Lloret, M.-R. (2005). More on Alignment as an Alternative to Domains: The Syllabification of Catalan Clitics. *Probus*, *17*, 37–78.

Bonet, E., Lloret, M.-R., & Mascaró, J. (2007). Allomorph Selection and Lexical Preferences: Two Case Studies. *Lingua*, *117*(6), 903–927.

Booij, G. (1998). Phonological output constraints in morphology. *Phonology and morphology of the Germanic languages*, *143*, 63.

Bradley, T. G. (2014). Optimality theory and spanish phonology. *Language and Linguistics Compass*, *8*(2), 65–88.

Broselow, E. (1992). Transfer and universals in second language epenthesis. In *Language transfer in language learning* (pp. 71–86). Amsterdam: John Benjamins.

Bye, P. (2007). Allomorphy – selection, not optimization. In S. Blaho, P. Bye, & M. Krämer (Eds.), *Freedom of analysis?* (pp. 63–92). Berlin: Walter de Gruyter.

Bye, P. & Svenonius, P. (2012). Non-concatenative morphology as epiphenomenon. In J. Trommer (Ed.), *The Phonology and Morphology of Exponence.* Oxford: Oxford University Press.

Caha, P. (2009). *The nanosyntax of case.* PhD thesis, Universitetet i Tromsø.

Caha, P. (2016). Notes on insertion in Distributed Morphology and Nanosyntax. *Ms. University of Tromsø, available at http://ling.auf.net/lingbuzz/002855.*

Caramazza, A. & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: evidence from 'the-tip-of-the-tongue' phenomenon. *Cognition, 64*(3), 309–343.

Cardinaletti, A. & Repetti, L. (2008). The Phonology and Syntax of Preverbal and Postverbal Subject Clitics in Northern Italian Dialects. *Linguistic Inquiry, 39*(4), 523–563.

Chomsky, N. (1995). *The minimalist program.* Cambridge, MA: MIT Press.

Chomsky, N. (1999). Derivation by phase. In *MIT Working Papers in Linguistics*, volume 18. Cambridge, MA: MIT Press.

Chomsky, N. & Halle, M. (1968). *The sound pattern of English.* New York: Harper & Row.

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. *Papers in laboratory phonology, 1*, 283–333.

Clua, E. (2002). Gènere i nombre en els noms i en els adjectius. In J. S. et al. (Ed.), *Gramàtica del català contemporani*, volume 1 (pp. 171–194). Barcelona: Empúries.

Clua, E., Goebl, H., Casassas, X., Civit, S., & Salicrú, M. (2013). Anàlisi dialectomètrica del cod amb el suport del vdm. In M.-R. Lloret & E. Clua (Eds.), *Questions de morfologia flexiva i lèxica del català. Volum d'homenatge a Joaquim Viaplana* (pp. 245–278). IIFV, Alacant.

Colina, S. (2003). The status of word-final [e] in Spanish. *Southwest journal of Linguistics*, *22*(1), 87.

Colina, S. (2007). Optimality-theoretic advances in our understanding of spanish syllable structure. In F. M. Gil & S. Colina (Eds.), *Optimality-theoretic studies in Spanish phonology* (pp. 278–311). Amsterdam: John Benjamins.

Corbett, G. G. (2000). *Number. Cambridge textbooks in linguistics.* Cambridge: Cambridge University Press.

Corbett, G. G. (2006). Gender, grammatical. *The Encyclopedia of Language and Linguistics. 2nd Edition. Oxford: Elsevier*, 749–756.

Coromines, J. (1936). *El parlar de Cardós i Vall Ferrera.* Barcelona: Casa de Caritat.

De Lacy, P. (2002). *The formal expression of markedness.* PhD thesis, University of Massachusetts Amherst.

del Gobbo, F. (2001). An ot account of the distribution of articles in italian. *Proceedings of SWOT*, *5*, 1–18.

Eguren, L. (1999). Pronombres y adverbios demostrativos. las relaciones deícticas. In I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española.* Madrid: Espasa Calpe.

Fabra, P. (1956). *Gramàtica Catalana.* Barcelona: Teide.

Fábregas, A. & Scalise, S. (2012). *Morphology: from data to theories.* Edinburgh: Edinburgh University Press.

Fleischhacker, H. (2001). Cluster-dependent epenthesis asymmetries. In A. Albright & T. Cho (Eds.), *UCLA working papers in linguistics*, volume 7 (pp. 77–116). UCLA.

Fleischhacker, H. A. (2005). *Similarity in phonology: Evidence from reduplication and loan adaptation.* PhD thesis, University of California Los Angeles.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology.* Cambridge, MA: MIT press.

Goldsmith, J. (1976). *Autosegmental Phonology.* PhD thesis, Massachusetts Institute of Technology.

Gouskova, M. (2002). Falling sonority onsets, loanwords, and syllable contact. In Andronis, M., Ball, C., Elston, H., & Neuvel, S. (Eds.), *CLS: The Main Session. Papers from the 37th Meeting of the Chicago Linguistic Society*, volume 1, (pp. 175–185).

Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. *Universals of language, 2*, 73–113.

Grimshaw, J. (1997). The best clitic: Constraint conflict in morphosyntax. In L. Haegeman (Ed.), *Elements of Grammar* (pp. 169–196). Dordrecht: Kluwer.

Hall, N. (2011). Vowel Epenthesis. In E. H. Marc Van Oostendorp, Colin J. Ewen & K. Rice (Eds.), *The Blackwell Companion to Phonology*. Oxford: Wiley-Blackwell.

Halle, M. (1997). Distributed morphology: Impoverishment and fission. In *MIT Working Papers in Linguistics*, volume 30 (pp. 425–449).

Halle, M. & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In K. Hale & S. J. Keyser (Eds.), *The View from Building 20* (pp. 111–176). Cambridge, MA: MIT Press.

Harley, H. & Noyer, R. (1999). Distributed morphology. *Glot international, 4*(4), 3–9.

Harley, H. & Ritter, E. (2002). Person and number in pronouns: A feature-geometric analysis. *Language, 78*(3), 482–526.

Harris, J. W. (1983). *Syllable Structure and Stress in Spanish*. Cambridge, MIT Press.

Harris, J. W. (1985). Spanish word markers. *Current issues in Hispanic phonology and morphology, 34*, 54.

Harris, J. W. (1991). The exponence of gender in spanish. *Linguistic Inquiry*, 27–62.

Harris, J. W. (1992). The form classes of spanish substantives, 65–88.

Harris, J. W. (1999). Nasal depalatalization no, morphological wellformedness sí; the structure of spanish word classes. *MIT Working Papers in Linguistics, 33*, 47–82.

Heap, D. (2005). Constraining optimality: Clitic sequences and feature geometry. In L. Heggie & F. Ordónez (Eds.), *Clitic and Affix Combinations: Theoretical Perspectives* (pp. 81–102). Amsterdam: John Benjamins.

Heath, J. (1998). Pragmatic skewing in pronominal 1 → 2 combinations in native american languages. *International Journal of American Linguistics*, 83–104.

Hinzen, W. & Sheehan, M. (2014). *The philosophy of universal grammar*. Oxford: Oxford University Press.

Hockett, C. F. (1954). Two models of grammatical description. *Word*, *10*(2-3), 210–234.

Jiménez, J. (2002). Altres fenòmens vocàlics en el mot. In J. S. et al. (Ed.), *Gramàtica del català contemporani*, volume 1 (pp. 171–194). Barcelona: Empúries.

Kayne, R. S. (2006). On parameters and on principles of pronunciation. In H. Broekhuis, N. Corver, R. Huybregts, U. Kleinhenz, & J. Koster (Eds.), *Organizing Grammar. Linguistic Studies in Honor of Henk van Riemsdijk* (pp. 289–299). Berlin: Mouton de Gruyter.

Kayne, R. S. (2008). Expletives, datives, and the tension between morphology and syntax. In T. Biberauer (Ed.), *The limits of syntactic variation* (pp. 175–217). Amsterdam: John Benjamins.

Kenstowicz, M. (1996). Base-identity and uniform exponence: alternatives to cyclicity. In J. Durand & B. Laks (Eds.), *Current Trends 17 in Phonology: Models and Methods*. European Studies Research Studies & University of Salford Press.

Kiparsky, P. (1982a). From cyclic phonology to lexical phonology. *The structure of phonological representations*, *1*, 131–175.

Kiparsky, P. (1982b). Lexical phonology and morphology. *Linguistics in the morning calm*.

Kiparsky, P. (1998). Paradigm effects and opacity. *Ms., Stanford University*.

Kiparsky, P. (2000). Opacity and cyclicity. *Linguistic review*, *17*(2/4), 351–366.

Kramer, R. (2009). *Definite markers, phi-features, and agreement: a morphosyntactic investigation of the Amharic DP*. PhD thesis, University of California Santa Cruz.

Kramer, R. (2014). Gender in Amharic: a morphosyntactic approach to natural and grammatical gender. *Language Sciences*, *43*, 102–115.

Kramer, R. (2015). *The morphosyntax of gender*. Oxford: Oxford University Press.

Labov, W. (1972). *Sociolinguistic patterns.* University of Pennsylvania Press.

Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language variation and change, 2*(2), 205–254.

Ladefoged, P. & Maddieson, I. (1996). *The sounds of the world1s languages.* Oxford: Blackwell.

Lampitelli, N. (2014). The Romance plural isogloss and linguistic change: a comparative study of Romance nouns. *Lingua, 140,* 158–179.

Leonetti, M. (1999). El artículo. In I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española.* Madrid: Espasa Calpe.

Levi, S. V. (2008). Phonemic vs. derived glides. *Lingua, 118*(12), 1956–1978.

Lieber, R. (1992). *Deconstructing morphology: Word formation in syntactic theory.* University of Chicago Press.

Lloret, M.-R. (1995). El tractament de les formes nominals' invariables' quant a nombre. *Caplletra: revista internacional de filología,* (19), 215–227.

Lloret, M.-R. (1998). Sobre l'estructura morfològica dels noms en català i en castellà. In G. Rufino (Ed.), *Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza (Università di Palermo, 18–24 settembre 1995)* (pp. 557–566). Niemeyer, Tubingen.

Lloret, M.-R. (2002). Estructura sil·làbica. In J. S. et al. (Ed.), *Gramàtica del català contemporani,* volume 1 (pp. 195–249). Barcelona: Empúries.

Lloret, M.-R. (2013). Diferències en les terminacions flexives dels elements nominals. In M.-R. Lloret & E. Clua (Eds.), *Questions de morfologia flexiva i lèxica del català. Volum d'homenatge a Joaquim Viaplana* (pp. 245–278). IIFV, Alacant.

Lloret, M.-R. (2015). La sufixació apreciativa del català: creacions lèxiques i implicacions morfològiques. *Caplletra. Revista Internacional de Filologia,* (58), 55–89.

Lloret, M. R. & Jiménez, J. (2008). Marcatge posicional i prominència en el vocalisme àton. *Caplletra. Revista Internacional de Filologia,* (45), 55–91.

Lloret, M.-R. & Jiménez, J. (2010). Sobre les vocals afegides de l´ alguerès. *INSULA: Quaderno di cultura sarda,* (7), 111–139.

Lloret, M.-R. & Pons-Moll, C. (2016). Catalan vowel epenthesis as evidence for the free ride approach to morphophonemic learning. *Linguistic Inquiry*, 147–157.

Lloret, M.-R. & Viaplana, J. (1992). Les vocals finals dels elements nominals en català. In A. Ferrando & A. G. Hauf (Eds.), *Miscel·lània Joan Fuster, vol. 5*. Abadia de Montserrat.

Lloret, M.-R. & Viaplana, J. (1997). On the morphological category of gender in Catalan and in Spanish. In W. U. Dressler, M. Prinzhorn, & J. R. Rennison (Eds.), *Advances in morphology*, volume 97 (pp. 171–188). Berlin: Mouton de Gruyter.

Longa, V. M., Lorenzo, G., & Rigau, G. (1998). Subject clitics and clitic recycling: locative sentences in some iberian romance languages. *Journal of linguistics*, *34*(01), 125–164.

Loporcaro, M. (1997). On vowel epenthesis in Alguer Catalan. In *Certamen Phonologicum III. Papers from the Third Cortona Phonology Meeting, April*, volume 199, (pp. 205–227).

Lowenstamm, J. (1996). Cv as the only syllable type. *Current trends in Phonology. Models and Methods, edited by Jacques Durand & Bernard Laks*, 419–441.

Marantz, A. (1997). No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania working papers in linguistics*, *4*(2), 14.

Marantz, A. (2001). Words. *WCCFL XX Handout, USC*.

Martín, F. J. (2012). *Deconstructing Catalan Object Clitics*. PhD thesis, New York University.

Martin, T. & Hinzen, W. (2014). The grammar of the essential indexical. *Lingua*, *148*, 95–117.

Mascaró, J. (1976). *Catalan phonology and the phonological cycle*. PhD thesis, Massachusetts Institute of Technology.

Mascaró, J. (1985). Catalan nominal markers and vowel insertion. *Cornell Working Papers in Linguistics*, *7*, 162–175.

Mascaró, J. (1986). Morfologia catalana. *Barcelona: Enciclopèdia Catalana*.

Mascaró, J. (1996). External allomorphy and contractions in romance. *Probus, 8*(2), 181–206.

Mascaró, J. (2007). External Allomorphy and Lexical Representation. *Linguistic Inquiry, 38*(4), 715–735.

Massanell, M. (2011). *Permeabilitat del català nord-occidental a l'estàndard mediàtic i educatiu: el cas de l'Alt Urgell.* Institut d'Estudis Catalans.

Massanell, M. (2012). *Feve temps que no diva tants verbs! Manteniment i transformación de paradigmas verbals en el català nord-occidental del tombant de segle.* Barcelona: Publicacions de l'abadia de Montserrat.

McCarthy, J. J. (1979). *Formal problems in Semitic phonology and morphology.* PhD thesis, Massachusetts Institute of Technology.

McCarthy, J. J. (1983). Consonantal morphology in the chaha verb. In *Proceedings of WCCFL2*, (pp. 176–188).

McCarthy, J. J. (1986). Ocp effects: Gemination and antigemination. *Linguistic inquiry, 17*(2), 207–263.

McCarthy, J. J. & Prince, A. (1993). Prosodic Morphology I: Constraint Interaction and Satisfaction. *Linguistics Department Faculty Publication Series* Paper 14.

McCarthy, J. J. & Prince, A. (1995). Faithfulness and reduplicative identity. In J. Beckman, L. Walsh Dickey, & Urbanczyk (Eds.), *University of Massachusetts Occasional Papers in Linguistics*, volume 18 (pp. 249–384).

Merchant, J. (2015). How much context is enough? Two cases of span-conditioned stem allomorphy. *Linguistic Inquiry.*

Morén, B. (2001). *Distinctiveness, coercion and sonority: A unified theory of weight.* Psychology Press.

Nevins, A. (2006). Dual is still more marked than plural. ms, Harvard University.

Nevins, A. (2011). Phonologically conditioned allomorph selection. In E. H. Marc Van Oostendorp, Colin J. Ewen & K. Rice (Eds.), *The Blackwell Companion to Phonology* (pp. 2357–2382). Oxford: Wiley-Blackwell.

Nevins, A. (2012). Morphophonological dissimilation, morphosyntactic dissimilation, and the architecture of exponence. In J. Trommer (Ed.), *The Handbook of Exponence.* Oxford: Oxford University Press.

Nevins, A. I. & Savescu, O. (2010). An apparent 'Number Case Constraint' in Romanian. In *Romance Linguistics 2008: Interactions in Romance. Selected papers from the 38th linguistic symposium on Romance languages (LSRL), Urbana-Champaign, April 2008*, volume 313, (pp. 185). Amsterdam: John Benjamins.

Newell, H. (2008). *Aspects of the Morphology and Phonology of Phases.* PhD thesis, McGill University.

Noyer, R. R. (1992). *Features, positions and affixes in autonomous morphological structure.* PhD thesis, Massachusetts Institute of Technology.

Oltra-Massuet, I. & Arregi, K. (2005). Stress-by-structure in spanish. *Linguistic Inquiry, 36*(1), 43–84.

Oltra-Massuet, M. I. (1999). *On the notion of theme vowel: A new approach to Catalan verbal morphology.* MA thesis, Massachusetts Institute of Technology.

Padgett, J. (2008). Glides, vowels, and features. *Lingua, 118*(12), 1937–1955.

Paster, M. E. (2006). *Phonological Conditions on Affixation.* PhD thesis, University of California.

Pena, J. (1999). Partes de la morfología. las unidades del anílisis morfolǿgico. In I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española.* Madrid: Espasa Calpe.

Pensado, C. (1999). Morfología y fonología. fenómenos morfofonológicos. In I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española.* Espasa Calpe.

Perea, M. P. (2002). Flexió verbal regular. In J. S. et al. (Ed.), *Gramàtica del català contemporani*, volume 1 (pp. 583–646). Barcelona: Empúries.

Pescarini, D. (2005). Clitic clusters and morphological repairs. evidence from Italian, Spanish and Barceloní. *Studi linguistici e filologici on line, 3*, 239–264.

Pescarini, D. (2010). Elsewhere in romance: Evidence from clitic clusters. *Linguistic Inquiry, 41*(3), 427–444.

Picallo, M. C. (1991). Nominals and nominalizations in catalan. *Probus, 3*(3), 279–316.

Picallo, M. C. (2006). Some notes on grammatical gender and l-pronouns. In von Heusinger, K., Kaiser, G. A., & Stark, E. (Eds.), *Proceedings of the workshop "Specificity and the evolution/emergence of nominal determination systems in Romance"*, (pp. 107–121).

Picallo, M. C. (2008). Gender and number in romance. *Lingue e linguaggio*, *7*(1), 47–66.

Pons-Moll, C. (2011). It is all downhill from here: A typological study of the role of syllable contact in romance languages. *Probus*, *23*(1), 105–173.

Pons-Moll, C. (2012). Loanword phonology, lexical exceptions, morphologically driven underapplication, and the nature of positionally biased constraints. In *Catalan journal of linguistics*, volume 11 (pp. 0127–166).

Pons-Moll, C. & Lloret, M.-R. (2014). *The free-ride procedure to morphophonemic learning is correct. Some evidence from Catalan vowel epenthesis.* Presented at the 22nd Manchester Phonology Meeting.

Prince, A. & Smolensky, P. (1993/2004). *Optimality Theory: Constraint interaction in generative grammar*. Oxford: Blackwell.

Repetti, L. (2006). Marked structures in loans in Italian. In *Amsterdam Studies in the Theory and History of Linguistic Science Series 4*, volume 274 (pp. 209–236). Amsterdam: John Benjamins.

Repetti, L. (2012). Consonant-final loanwords and epenthetic vowels in Italian. *Catalan journal of linguistics*, *11*, 0167–188.

Roca, I. (2010). Theme vowel allomorphy in Spanish verb inflection: An autosegmental optimality account. *Lingua*, *120*(2), 408–434.

Rosenthall, S. (1997). *Vowel/glide alternation in a theory of constraint interaction.* PhD thesis, University of Massachusetts-Amherst.

Sayers, D. (2014). The mediated innovation model: A framework for researching media influence in language change. *Journal of Sociolinguistics*, *18*(2), 185–212.

Scalise, S. (1983). *Morfologia lessicale.* Padova: CLESP.

Scheer, T. (2012). *Direct interface and one-channel translation*, volume 68. Berlin: Walter de Gruyter.

Siddiqi, D. (2009). *Syntax within the word: Economy, allomorphy, and argument selection in Distributed Morphology*, volume Linguistik Aktuell 138. Amsterdam: John Benjamins.

Sistac, R. (1993). *El ribagorçà a l'alta Llitera: els parlars de la vall de la Sosa de Peralta.* Institut d'Estudis Catalans.

Sistac, R. (1998). *El Català d'Àneu: reflexions a l'entorn dels dialectes contemporanis.* Consell Cultural de les Valls d'Àneu.

Starke, M. (2009). Nanosyntax: A short primer to a new approach to language. *Nordlyd, 36*(1), 1–6.

Steriade, D. (1997). Phonetics in phonology: the case of laryngeal neutralization. *UCLA, manuscript.*

Steriade, D. (1999). Lexical conservatism in french adjectival liaison. In B. Bullock, M. Authier, & L. Reed (Eds.), *Formal perspectives in Romance linguistics* (pp. 243–270). Philadelphia: John Benjamins.

Stump, G. T. (2001). *Inflectional Morphology: A theory of Paradigm Structure.* Cambridge: Cambridge University Press.

Suïls, J. (1993). *Els clítics pronominals de tercera persona al subdialecte ribagorçà.* PhD thesis, MA thesis, Universitat Autònoma de Barcelona.

Svenonius, P. (2012). Spanning. *Ms. University of Tromsø, available at ling.auf.net/lingBuzz/001501.*

Todolí, J. (2002). Els pronoms. In J. e. a. Solà (Ed.), *Gramàtica del català contemporani*, volume 2 (pp. 1337–1433). Barcelona: Empúries.

Tranel, B. (1996). Exceptionality in optimality theory and final consonants in french. In K. Zagona (Ed.), *Grammatical Theory and Romance Languages* (pp. 275–293). Amsterdam: John Benjamins.

Trommer, J. (2002). *Distributed optimality.* PhD thesis, Universität Potsdam.

Trommer, J. (2011). *Phonological aspects of Western Nilotic mutation morphology.* PhD thesis, Universität Leipzig.

Trommer, J. (2012). Ø exponence. In J. Trommer (Ed.), *The morphology and phonology of exponence*, volume 41. Oxford: Oxford University Press.

Trommer, J. & Zimmermann, E. (2015). Inflectional exponence. In M. Baerman (Ed.), *The Oxford Handbook of Inflection*. Oxford University Press.

Trudgill, P. (2014). Diffusion, drift, and the irrelevance of media influence. *Journal of Sociolinguistics*, *18*(2), 213–222.

Valls, E. (2013). Linguistic advergence and divergence in northwestern Catalan: A dialectometric investigation of dialect leveling and border effects. *Literary and Linguistic Computing*, 119–146.

Vallverdú, T. (2002). Fenòmens en grups vocàlics. In J. S. et al. (Ed.), *Gramàtica del català contemporani*, volume 1 (pp. 171–194). Barcelona: Empúries.

van Oostendorp, M. (2005). Expressing inflection tonally. *Catalan Journal of Linguistics*, *4*, 107–127.

van Oostendorp, M. (2015). Where syntax and phonology touch. a representational view of the interface. *12th Old World Conference in Phonology, Barcelona, 20-30th January*.

Vennemann, T. (1988). *Preference Laws for Syllable Structure: And the Explanation of Sound Change with Special Reference to German, Germanic, Italian, and Latin*. Walter de Gruyter.

Viaplana, J. (1980). Algunes consideracions sobre les formes pronominals clítiques del barceloní. *Anuario de filología*, (6), 459–484.

Viaplana, J. (1986). Morfologia flexiva i flexió verbal catalana. *Llengua & Literatura, 1986, vol. 1, p. 385-403*.

Viaplana, J. (1991). Sobre la flexió nominal en català. In *Estudis de Llengua i Literatura Catalanes, vol. 23 (M. Jordi Carbonell, 2)* (pp. 235–264). Publicacions de l'Abadia de Montserrat.

Viaplana, J. (1999). *Entre la dialectologia i la linguística. La distància linguística entre les varietats del català nordoccidental*. Publicacions de l'Abadia de Montserrat.

Viaplana, J., Lloret, M.-R., Perea, M.-P., & Clua, E. (2007). *COD. Corpus Oral Dialectal*. Barcelona: PPU.

Vigário, M. (1999). Pronominal cliticization in European Portuguese. *Catalan working papers in linguistics*, *7*, 219–237.

Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of italian tongues. *Psychological science*, *8*(4), 314–317.

Wheeler, M. (1979). *Phonology of Catalan.* Oxford: Basil Blackwell.

Wheeler, M. (2005). *The Phonology of Catalan.* Oxford: Oxford University Press.

Xavier, C. (2013). Is the production of epenthesis in loanwords influenced by English proficiency? A production study in Brazilian Portuguese. *3rd International Conference on English Pronunciation, 8-10th May, Murcia.*

Yun, S. (2012). Perceptual similarity and epenthesis positioning in loan adaptation. In *Proceedings of the 48th Annual Meeting of Chicago Linguistics Society (CLS 48).*

Yun, S. (2014). The role of acoustic cues in nonnative cluster repairs. In Santana-LaBarge, R. E. (Ed.), *Proceedings of the 31st West Coast Conference on Formal Linguistics (WCCFL 31)*, (pp. 514–523). Somerville, MA: Cascadilla Proceedings Project.