The structure of allomorphy in Spanish verbal inflection
Gilles Boyé & Patricia Cabredo Hofherr

Introduction

In the present paper we propose an analysis of Spanish verbal inflection which accounts for the limits on allomorphy observed in the paradigms. Some of these restrictions on allomorphy are well-known (Bello 1970, chapter XXIV, §503-§504). For example, in Spanish, every verb has a single stem for all Future Indicative and Conditional forms, and for all verbs, the stem of the infinitive and of the 2nd plural Imperative are identical (amaveramad; ser—sed).

In section 1, we present the morphological analysis of verbal forms that we will assume for the purpose of our analysis of suppletion patterns.

In section 2, we will show, following work by Aronoff and Pirrelli & Battista, that this observation can be generalised and that the verbal paradigm can be partitioned into groups of verbal forms that always pattern together even in the most irregular verbs.

In section 3, we will examine the distribution of irregularities across the stems established in section 2. Since the observed irregularities do only exploit a fraction of the possible patterns, we will propose, that the stem space is itself structured. Taking Boyé (2000, chapter 6) as a starting point, we use the work of Bonami & Boyé (2002, 2003) on French to further refine the analysis of the restrictions on stem allomorphy in Spanish conjugation. More specifically, we follow Bonami & Boyé (2003) and develop an analysis that relies on implicational relations between the stems which limit the distribution of stem allomorphy.

1. The morphological analysis of verbal forms

Traditionally, Spanish verbal forms have been analysed into four parts (see e.g. Alcoba 1999: 4924, and references cited there):

(1)  
   a. radical + thematic vowel + Tense/Aspect/Mode + Person/Number  
   b. cant + a + b a+ s  
      ‘you sang’ (imperfective)

The thematic vowel is taken to indicate the conjugation class a verb belongs to. However, it is well-known that irregular verbs can have different thematic vowels in different parts of their paradigm. For example, andar behaves like a verb with a thematic -a- (amar) except for the Preterite and the Imperfective and Future Subjunctive where its forms display a thematic -i- (anduv-i-mos vs. am-a-mos).

(2)  
   amar  ‘to love’  
   am+a+s  
   am+a+ ste  
   andar  ‘to walk’  
   and+a+s  
   anduv+i+ste  
   2sg Present Indicative  
   2sg Preterit

This suggests that the thematic vowel (TV) is not a property of the verb as a whole but rather a property of each of the stems (/anda/ and /anduvi/ in the case of andar), as noted by Aronoff (1994, pXXX) for Latin conjugation.

---

1 We are grateful to M. Plénat for providing us with this reference.
We will assume a morphological analysis of the verbal form into two parts: a stem and an affix which for simplicity we will consider to be unanalysed.²

(3) a. stem + ending
   b. cantad + o

The decomposition into two parts is based on the possible suppletions:

(4) Hypothesis 1: The ending is the invariant suffix that appears with suppletive stems.

Given this hypothesis, the suffix of the past participle is -o:

(5) a. cantad + o
    b. dich + o

Following Hypothesis 1 strictly, several forms would have to be analysed as having a zero suffix [Ø] as in the following two examples:

(6) 3sg Present Indicative
    a. (él) ama + Ø
    b. (él) come + Ø
    c. (él) vive + Ø

(7) 1sg Present Indicative (first version)
    a. (yo) amo + Ø
    b. (yo) sé + Ø
    c. (yo) he + Ø

The two examples are clearly different though: the first analysis is based on a robust contrast that can be found across all verbs, furthermore the stem postulated on the basis of (6) resurfaces in the 2sg (ama-s, come-s, vive-s) and the 3pl (ama-n, come-n, vive-n) of the Present Indicative. The contrast in (7), on the other hand, is based on few forms only (sé, he, soy, estoy, voy, doy) which are clearly unanalysable exceptions to the regular inflection pattern, and the putative stem /sé/ would not appear anywhere else in the paradigm of saber. We will therefore assume the following Hypothesis 2:

(8) Hypothesis 2: The verbal forms which are idiosyncratic with respect to their endings are excluded from the comparisons establishing the invariant suffix.

Applying Hypothesis 2 excludes (7b/c) from the comparison and gives the following morphological analysis for the 1sg Present Indicative:

(9) 1sg Present Indicative (second version)
    a. (yo) am + o
    b. (yo) com + o
    c. (yo) viv + o

² In section 3, we adopt a realizational model of inflectional morphology within which the ending is only a part of the function realizing the inflected form.
Notice that, following Bonami & Boyé (2002), we distinguish irregular but analysable forms ((yo) caig+o) from irregular non-analysable forms ((yo) sé). We will call suppletions of the caigo-type STEM-SUPPLETION, since only the irregular stem has to be lexically specified. In contrast to this, suppletions of the sé-type are suppletions of the entire inflected form; we will call this type of suppletion FORM-SUPPLETION in what follows.

(10) a. stem + ending
b. caig + o (stem suppletion, analysable)
c. sé (form-suppletion, unanalysable)

Following this approach we arrive at the following analysis of the verbal forms:

(11)

<table>
<thead>
<tr>
<th></th>
<th>1sg</th>
<th>2sg</th>
<th>3sg</th>
<th>1pl</th>
<th>2pl</th>
<th>3pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present Ind.</td>
<td>±o</td>
<td>+s</td>
<td>+Ø</td>
<td>+mos</td>
<td>+js</td>
<td>+n</td>
</tr>
<tr>
<td>Present Subj.</td>
<td>+Ø</td>
<td>+s</td>
<td>+Ø</td>
<td>+mos</td>
<td>+js</td>
<td>+n</td>
</tr>
<tr>
<td>Imperfective Ind.</td>
<td>+a</td>
<td>+as</td>
<td>+a</td>
<td>+amos</td>
<td>+ajs</td>
<td>+an</td>
</tr>
<tr>
<td>Preterite</td>
<td>±Ø</td>
<td>+ste</td>
<td>±o</td>
<td>+mos</td>
<td>+stejs</td>
<td>±ron</td>
</tr>
<tr>
<td>Imperfective 1 Subj.</td>
<td>±ra</td>
<td>±ras</td>
<td>±ra</td>
<td>±ramos</td>
<td>±rajs</td>
<td>±ran</td>
</tr>
<tr>
<td>Imperfective 2 Subj.</td>
<td>±se</td>
<td>±ses</td>
<td>±se</td>
<td>±semos</td>
<td>±sejs</td>
<td>±sen</td>
</tr>
<tr>
<td>Future Subj.</td>
<td>±re</td>
<td>±res</td>
<td>±re</td>
<td>±remos</td>
<td>±rejs</td>
<td>±ren</td>
</tr>
<tr>
<td>Future Ind.</td>
<td>+re</td>
<td>+ras</td>
<td>+ra</td>
<td>+remos</td>
<td>+rejs</td>
<td>+ran</td>
</tr>
<tr>
<td>Conditional</td>
<td>+rija</td>
<td>+rijas</td>
<td>+rija</td>
<td>+rijamos</td>
<td>+rijajs</td>
<td>+rijan</td>
</tr>
<tr>
<td>Imperative</td>
<td>+Ø</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infinitive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present participle</td>
<td></td>
<td>+ndo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past participle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+o</td>
</tr>
</tbody>
</table>

Notice that in the Imperfective the segments -áb- and -í- are analysed as part of the stem. This is motivated by the fact that the endings postulated are found in forms with suppletive stems such as ér-a and íb-a:

(12) a. (yo) amab + a 1sg Preterit ‘love’
b. (yo) comí + a 1sg Preterit ‘eat’
c. (yo) er + a 1sg Preterit ‘be’ (*er-ía, *er-aba)
d. (yo) ib + a 1sg Preterit ‘go’ (*ib-ía, *ib-aba)

Given the suffixes that we propose in the table in (11), we consider the following forms to be form-suppletions:

(13) Form-suppletions
1sg Present Indicative soy, estoy, voy, doy, hé, sé, hé³
3sg Preterit Indicative fue

In the table in (11), we have marked the endings for which the analysis is problematic by ±. We will come back to these forms in section 2, when we discuss the partition into stems.

³ The 2sg Imperative form of haber is not commonly used but Spanish native speakers still have the intuition that the form is hé.
2. The verb and its stems

Given a morphological analysis of Spanish verbal inflection into stem+ending, it can be observed that irregular forms do not fill arbitrary sets of cells in the paradigm, but rather cover parts of the paradigm that systematically use the same allomorph. Well-known examples of forms which always share the same stem are:

(14) a. the Future Indicative and Conditional forms 
\( (ama+ré—ama+ría; tend+ré—tend+ría) \),
b. the Infinitive and the 2nd plural Imperative
\( (ama+r—ama+d; se+r—se+d) \).

Pirrelli & Battista (2000) propose an analysis of Italian verb inflection that extends this type of generalisation to obtain a partition of the entire paradigm into areas which always share the same stem. This partition limits the possible stem suppletions: if one form of an area is based on a suppletive stem, every other form of the same area has to be based on the same suppletive stem. Following Bonami & Boyé (2003), we will call the list of stems defined by this partition the STEM-SPACE. The stem-space formalises the insight of traditional Latin grammars that the forms of an irregular verb can all be predicted based on the principal parts of the verb (which traditionally are given in the form of a list: perdo-perdere-perdidi-perditum, ‘lose’).

As pointed out by Aronoff (1994), the forms which systematically share an allomorph do not always form a natural class semantically, syntactically or phonologically, i.e. the allomorph cannot be characterised as expressing a particular set of features. This is illustrated by the example of the infinitive and the 2nd plural imperative in Spanish. Following Aronoff, we assume that the patterning into stems is purely morphological, and that there is a morphological organisation of the paradigms. In the case at hand the infinitive and the 2nd plural Imperative form a natural class only within the morphological organisation. The stem that serves as a base for both these two forms is a MОРPHOME in Aronoff’s sense. The hypothesis that there is a morphological organisation of the verbal inflection is supported independently by the observations in Maiden (1992), who points out that diachronic changes in the inflection of Italian occurred along the lines of such a partition, and not only in ways motivated by a phonological or a syntactic feature values.

In what follows, we develop an analysis for Spanish verbal inflection along the lines of Pirrelli & Battista (2000).

Taking the possibility of isolated suppletion as a criterion for distinct stems we arrive at the following partition:\(^4\):

(15)

<table>
<thead>
<tr>
<th></th>
<th>1sg</th>
<th>2sg</th>
<th>3sg</th>
<th>1pl</th>
<th>2pl</th>
<th>3pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>( -o )</td>
<td>+s</td>
<td>+( \emptyset )</td>
<td>+mos</td>
<td>+js</td>
<td>+n</td>
</tr>
</tbody>
</table>

\(^4\) In this table we exclude the forms which we have independently analysed as form-suppletions in section 1, and which therefore by definition do not have a stem+ending structure. The partition is represented by the bold lines. The grey cells present stem allomorphies within suppletion zones (see the discussion in section 3).
The table (15) distinguishes 13 zones:

(16)  
1sg Present Indicative and Present Subjunctive  
2sg Present Indicative  
3sg Present Indicative  
1pl Present Indicative  
2pl Present Indicative  
3pl Present Indicative  
Preterit / Imperfective 1&2 Subjunctive and Future Subjunctive  
Imperfective Indicative  
Future Indicative and Conditional  
2sg Imperative  
Infinitive and 2pl Imperative  
Present participle  
Past participle

The maximal stem-space gives the maximally possible number of stem-suppletions and therefore treats all forms as analysable, if only with a zero affix +Ø. In what follows we will discuss whether some irregular forms are better analysed as form-suppletions. Form-suppletions are stored as unanalysable units which do not have a stem and therefore they do not enter into the generalisations concerning the organisation of the stem-space.

Some zones in (16) cover the classical observations in (14) above, other stems are distinguished by an idiosyncratic stem-suppletion which combines with a non-zero suffix:

The table (15) distinguishes 13 zones:

<table>
<thead>
<tr>
<th>Present</th>
<th>Imperfective</th>
<th>Preterite</th>
<th>Imperfective</th>
<th>Future</th>
<th>Conditional</th>
<th>Imperative</th>
<th>Present participle</th>
<th>Past participle</th>
</tr>
</thead>
<tbody>
<tr>
<td>+Ø</td>
<td>+a</td>
<td>±Ø</td>
<td>±ra</td>
<td>±re</td>
<td>+rija</td>
<td>+Ø</td>
<td>±ndo</td>
<td>±o</td>
</tr>
<tr>
<td>+s</td>
<td>+as</td>
<td>+ste</td>
<td>±ras</td>
<td>±res</td>
<td>+rijas</td>
<td>+r</td>
<td>+t</td>
<td></td>
</tr>
<tr>
<td>+Ø</td>
<td>+a</td>
<td>±o</td>
<td>±ra</td>
<td>±ra</td>
<td>+rajs</td>
<td>+mos</td>
<td>+anim</td>
<td></td>
</tr>
<tr>
<td>+mos</td>
<td>+amos</td>
<td>+amos</td>
<td>±ramos</td>
<td>±rams</td>
<td>+ran</td>
<td>+js</td>
<td>+an</td>
<td></td>
</tr>
<tr>
<td>+js</td>
<td>+ajs</td>
<td>+stejs</td>
<td>±rajs</td>
<td>±rejs</td>
<td>+rjan</td>
<td>+n</td>
<td>+a</td>
<td></td>
</tr>
<tr>
<td>+n</td>
<td>+an</td>
<td>±ron</td>
<td>±ran</td>
<td>±ren</td>
<td></td>
<td>+d</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(17)  
a. Past participle  
   dich+o, muert+o, puest+o, vuelt+o, abiert+o  
b. Imperfective Indicative  
   er+a, ib+a  
c. 1pl Present Indicative  
   he+mos, so+mos  
d. 2pl Present Indicative  
   so+is  
e. 3 pl Present Indicative  
   ha+n, so+n

Other zones are motivated only by forms that have a zero suffix +Ø. If a form has a non-zero suffix (as for the Past participle), we postulate a cell in the stem-space, even if it is used only for a single cell of the paradigm. If the putative suffix is zero (+Ø), however, the question arises whether the forms under scrutiny are better analysed as form-suppletions. We will opt
for the analysis by form-suppletion if only a few verbs motivate a separate stem with a zero affix as for the following examples:

(18)  a. 3sg Present Indicative always shares the stem of 2sg Present Indicative except for ser: (él) es vs. (tú) ere+s.
    b. 2sg Imperative and 3sg Present Indicative are always identical, apart from 8 exceptional monosyllabic forms haz, ten, ven, pon, di, sal, sé, ve.

According to our analysis, the regular 2sg Imperative and 3sg Present Indicative have a zero suffix +Ø; it is therefore impossible to decide on the basis of the surface form if the forms motivating the distinctions 3sg Present Indicative vs. 2sg Present Indicative and 3sg Present Indicative vs. 2sg Imperative are to be analysed as stem-suppletions with a zero suffix (es+Ø, haz+Ø) or rather as form-suppletions (es, haz). If we treat these cases as stem-suppletions, we have to assume a separate 2sg Imperative stem for haz, ten, ven, pon, di, sal, sé, ve and a separate 3sg Present Indicative stem es. These stems, however, would be used only for the 2sg Imperative and the 3sg Present Indicative, respectively. It therefore seems preferable to analyse these cases as form-suppletions which are stored directly and do not participate in the organisation of the stem-space. Consequently, we assume a single stem for 2sg Present Indicative, 3sg Present Indicative and the 2sg Imperative.

The preceding discussion reduces the stem-space to the following 11 stems:

<table>
<thead>
<tr>
<th>Stem Description</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg Present Indicative and Present Subjunctive</td>
<td>S1</td>
</tr>
<tr>
<td>2-3sg Present Indicative and 2sg Imperative</td>
<td>S2</td>
</tr>
<tr>
<td>1pl Present Indicative</td>
<td>S3</td>
</tr>
<tr>
<td>2pl Present Indicative</td>
<td>S4</td>
</tr>
<tr>
<td>3pl Present Indicative</td>
<td>S5</td>
</tr>
<tr>
<td>Preterit (pretérito), Imperfective 1&amp;2 Subjunctive and Future Subjunctive</td>
<td>S6</td>
</tr>
<tr>
<td>Imperfective Indicative</td>
<td>S7</td>
</tr>
<tr>
<td>Future Indicative and Conditional</td>
<td>S8</td>
</tr>
<tr>
<td>Infinitive and 2pl Imperative</td>
<td>S9</td>
</tr>
<tr>
<td>Present participle</td>
<td>S10</td>
</tr>
<tr>
<td>Past participle</td>
<td>S11</td>
</tr>
</tbody>
</table>

The table in (19) gives the stems that can be suppled independently. However, for the stems S1 and S6, the suppletion zones identified by the table (19) do not coincide with a single surface stem in the morphological analysis in table (15).

The problem that arises is the following: even for regular verbs, the morphological analysis assumed here gives two different stem-forms for the zone covering Present Subjunctive and the 1sg Present Indicative and four different stem-forms for the inflected forms in the zone Preterite, Imperfective 1&2 Subjunctive and Future Subjunctive.

We will look at the two cases in turn. First, consider the following forms of the Present Subjunctive and the 1sg Present Indicative:

<table>
<thead>
<tr>
<th>Stem Description</th>
<th>Present Subjunctive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg Present Indicative</td>
<td>Present Subjunctive</td>
</tr>
<tr>
<td>a. am</td>
<td>a. am</td>
</tr>
<tr>
<td></td>
<td>+o</td>
</tr>
<tr>
<td>b. beb</td>
<td>beba +Ø</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
As the examples show, the stem that appears in the 1sg Present Indicative systematically differs from the stem that surfaces in the Present Subjunctive. Notice, however, that even for suppletive verbs the relationship between the two stems is completely regular: the 1sg Present Indicative stem corresponds to the Present Subjunctive stem after truncation of the thematic vowel.

The distinction between these two stems is not reflected in any stem-suppletion pattern. If the stem-space based on suppletion patterns captures all the stems visible for the morphological organisation, the two stems that surface in (20) have to receive a different status. In what follows, we propose an analysis that is based on the following hypothesis

(21) If a set of forms always function together with respect to suppletion they share a common underlying stem.

This hypothesis excludes that two stems may have systematically related values while occupying two positions in the stem-space. This then implies that the forms in (20) rely on a single stem (S1: *ame, caiga*) while displaying different surface stems (S1a: *am, caig* vs. S1b: *ame, caiga*). To maintain this hypothesis, we adopt a model of morphology where a form is realized by a function applied to the stem (see for example Stump 2001) rather than by the concatenation of the stem and an affix. This analysis is realizational: the functions do not add features to the stem but only express a feature bundle already present in the representation of the inflected form they are realizing.

We propose that, in the case at hand, the functions realizing the forms can be decomposed into two sub-functions: the first one (FS) creates the surface-stems that appear in (20) and second one (F+A) concatenates the surface-stem with the suffixes. The forms are morphologically derived from a single underlying stem as in the schema in (22) through the combination of the sub-functions.

(22) Underlying Stem => Surface Stem => Inflected Form
  a.  FS: Underlying Stem => Surface Stem
  b.  F+A: Surface Stem => Surface Stem + A = Inflected Form
  c.  F+A(FS(Underlying Stem)) = Inflected Form

For the stem S1, we have two surface stems S1a and S1b respectively for 1sg Present Indicative and for Present Subjunctive. In (23) and (24), we give the functions $F_{S1a}$ and $F_{S1b}$ realizing the two surface stems from the underlying ones.

(23) $F_{S1a}$: S1 => Surface Stem for 1sg Present Indicative: S1a
     S1 --------------> S1a
     truncate TV$^5$

5 See the beginning of section 1, where we have argued that the thematic vowel is local to a stem.
a. $F_{S1a}(ame) = am$
   $F_{S1a}(beba) = beb$
   $F_{S1a}(viva) = viv$

b. $F_{S1a}(caiga) = caig$
   $F_{S1a}(quepa) = quep$
   $F_{S1a}(diga) = dig$

(24) $F_{Sib}$:  
S1 => Surface Stem for Present Subjunctive: S1b
S1 ---------> S1b
   identity

  a. $F_{Sib}(ame) = ame$
  $F_{Sib}(beba) = beba$
  $F_{Sib}(viva) = viva$

  b. $F_{Sib}(caiga) = caiga$
  $F_{Sib}(quepa) = quepa$
  $F_{Sib}(diga) = diga$

We can now define the functions $F_n$ realizing the different inflected forms based on S1 as the combination of two sub-functions:

(25) a. $F_1(S1) => 1$sg Present Indicative  e.g. $F_1(ame)=amo$

b. $F_2(S1) => 1$sg Present Subjunctive  e.g. $F_2(ame)=ame$

c. $F_3(S1) => 2$sg Present Subjunctive  e.g. $F_3(ame)=ames$

where $F_1(S1) = F_{eo}(F_{S1a}(S1)) = (S1-VT)+o$
ame -------------------------> am+o
   truncate TV, add +o

$F_2(S1) = F_{eÔ}(F_{Sib}(S1)) = (S1)+Ô$
ame -------------------------> ame+Ô
   keep the stem, add +Ô

$F_3(S1) = F_{es}(F_{Sib}(S1)) = (S1)+s$
ame -------------------------> ame+s
   keep the stem, add +s

Now consider the four different surface stems for the inflected forms in the suppletion zone Preterite, Imperfective 1&2 Subjunctive and Future Subjunctive (S6):

(26) regular          suppletive
    1sg Preterite       amé +Ô  bebí +Ô  tuve +Ô
S6a 1sg Preterite       amé +Ô  bebí +Ô  tuve +Ô
S6b 2sg, 1-2pl Preterite amá +ste bebí +ste tuvi +ste
S6c 3sg Preterite       am +ó  bebí +ó  tuv +o
S6d 3pl Preterite, Imperf. 1&2 and Fut. Subj. amá +ron bebié +ron tuvié +ron

Again, the distinctions between these four surface stems is not reflected by the existence of verbs that are clearly suppletive in this zone; the only possible suppletions are change or truncation of the thematic vowel, change of accent, diphthongisation of the thematic vowel. Notice that these changes are also present in the irregular pattern, exemplified by 
tener above:
the surface variations in the stems of the Preterit are the same for regular and irregular verbs. The three patterns in (26) are the only patterns of stem variation possible in the Preterit.\(^6\)

Like for the relation between the surface stems and the underlying stem in the Present Subjunctive and the 1sg Present Indicative, we propose that the surface-stems in the cases above are morphologically conditioned by the thematic vowel of the underlying stem. There are three possibilities for the underlying thematic vowel: a stressed /á/ (amar: amá), a stressed /í/ (beber: bebí) and unstressed /i/ (tener: tuvi). The functions deriving the four surface stems are given below in (27–30).

(27) S6 => Surface Stem for 1sg Preterit: S6a  
  F\(_{S6a}\): S6 => S6a = S6 with a modification of the TV: á => é, í => í, i => e  
  S6a:  
  \(F_{S6a}(amá) = amé\)  
  \(F_{S6a}(bebí) = bebí\)  
  \(F_{S6a}(tuvi) = tuve\)

(28) S6 => Surface Stem for 2sg, 1-2pl Preterit: S6b  
  F\(_{S6b}\): S6 => S6b = S6 with a modification of the TV: á => á, í => í, i => í  
  S6b:  
  \(F_{S6b}(amá) = amá\)  
  \(F_{S6b}(bebí) = bebí\)  
  \(F_{S6b}(tuvi) = tuví\)

(29) S6 => Surface Stem for 3sg Preterit: S6c  
  F\(_{S6c}\): S6 => S6c = S6 with a modification of the TV: á => Ø, í => i, i => Ø  
  S6c:  
  \(F_{S6c}(amá) = am\)  
  \(F_{S6c}(bebí) = bebi\)  
  \(F_{S6c}(tuvi) = tuv\)

(30) S6 => Surface Stem for 3pl Preterit / Imperf 1&2 and Future Subjunctive: S6d  
  F\(_{S6d}\): S6 => S6d = S6 with a modification of the TV: á => á, í => ié, i => ié  
  S6d:  
  \(F_{S6d}(amá) = amá\)  
  \(F_{S6d}(bebí) = bebié\)  
  \(F_{S6d}(tuvi) = tuvié\)

This analysis assumes regular variations of the surface-stem conditioned by the thematic vowel.\(^7\) Notice that this differs from the traditional analysis only insofar as the traditional analysis postulates a global thematic vowel for a verb which conditions arbitrary changes in thematic vowel (for example Present Indicative \(am-a-s\) as opposed to Present Subjunctive \(am-e-s\)), or in the choice of affix (for example \(-aba\) as opposed to \(-a\)), whereas we postulate a local thematic vowel for a stem which conditions arbitrary changes in the stem itself.\(^8\)

---

\(^6\) The forms of the Imperfective and Future Subjunctive always share the stem of the 3pl Preterit Indicative.

\(^7\) The verbs in class 2 that show an alternation between \(-zc-\) and \(-c-\) (parezco vs parece) can be considered to be regular if we assume that their underlying stems contain \(-zc-\) throughout. If followed by \(-a-\), \(-o-\) or \(-u-\), the final \(-c-\) surfaces as [k], if followed by \(-e-\) or \(-i-\), the final \(-c-\) surfaces as [0] or [s] (see Boyé 2000, pp 260–261).

\(^8\) For an analysis which gives a uniform morphophonological analysis of the variation of the surface stem in (26), see Boyé (2001).
Summarising, we have argued that the paradigm of Spanish verbal inflection can be partitioned into 11 suppletion zones. We have argued that the limited surface variation that is observable within the suppletion zone corresponding to the Preterite / Imperfective 1&2 and Future Subjunctive is not suppletion but regular variation conditioned by the thematic vowel of the underlying stem. Given our hypothesis that the only stems relevant for the morphological organisation are the stems identified by suppletion, the regular variation of the surface stem as in (20) and (26) is part of the function that constructs the inflected forms on the basis of a stem.

Within the stem-space, each stem represents a zone that may be suppleted separately. However, it can be observed that suppletions do not target all the possible subsets of the stem-space. In the following section we explore the hypothesis proposed for French by Bonami & Boyé (2002, 2003) that the stem-space is structured by local relations between the stems which restrict the possible suppletion patterns.

3. Stem relations

The stem-space that we have argued for in section 2 groups forms together which are always identically suppleted for all morphologically analysable verbs. Regular verbs only specify one stem lexically, the other stems are derived through relations from this stem. The relations between stems are defined on the basis of the regular verbs. Suppletive stems lexicalise a form distinct from the expected stem. Given the morphological analysis presented in section 1, the stems of a regular verb are linked by these relations, this does not entail, however, that the stems are identical.

Take the following examples:

<table>
<thead>
<tr>
<th>verb</th>
<th>Pres.Subj.</th>
<th>1sg Pres. Ind.</th>
<th>2/3 sg Pres. Ind.</th>
<th>2sg Imperative</th>
<th>1pl Pres. Ind.</th>
<th>2pl Pres. Ind.</th>
<th>3pl Pres. Ind.</th>
<th>Preterit, Imperf 1&amp;2 Fut Subj</th>
<th>Imperfect</th>
<th>Future</th>
<th>Conditional</th>
<th>Infinitive</th>
<th>2pl Imperative</th>
<th>Present</th>
<th>Participle</th>
<th>Past Participle</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMAR</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>ame</td>
<td>ama</td>
<td>ama</td>
<td>ama</td>
<td>ama</td>
<td>ama</td>
<td>amá</td>
<td>amá</td>
<td>amá</td>
<td>amá</td>
<td>amá</td>
<td>amá</td>
<td>amá</td>
<td>amad</td>
<td>amad</td>
<td>amad</td>
</tr>
<tr>
<td>BEBER</td>
<td>A</td>
<td>beba</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>beba</td>
<td>bebe</td>
<td>bebe</td>
<td>bebe</td>
<td>bebe</td>
<td>bebe</td>
<td>bébí</td>
<td>bébí</td>
<td>bébí</td>
<td>bébí</td>
<td>bébí</td>
<td>bébí</td>
<td>bébí</td>
<td>bébíd</td>
<td>bébíd</td>
<td>bébíd</td>
</tr>
<tr>
<td>VIVIR</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>viva</td>
<td>vive</td>
<td>viví</td>
<td>vivi</td>
<td>viví</td>
<td>viví</td>
<td>viví</td>
<td>viví</td>
<td>viví</td>
<td>viví</td>
<td>viví</td>
<td>vivió</td>
<td>viví</td>
<td>vici</td>
<td>vici</td>
<td>vici</td>
</tr>
</tbody>
</table>

The alternations observed in (31) are the regular stem changes linking the different stems of regular verbs. In what follows we will call two stems that stand in the regular allomorphic relation to one another RELATED STEMS.

\[9\] We therefore distinguish two types of allomorphy:

(i) stem changes pertaining to the paradigm of regular verbs (regular stem changes)

(ii) different stems independently lexicalised (suppletion and irregular stem changes).
Following the traditional analysis, we will assume that the thematic vowel conditions regular alternations within the regular paradigm. For regular verbs, the possible variations between allomorphic stems are located at the site of the thematic vowel.

Taking for example the thematic vowel of the Infinitive/2pl Imperative stem (S9) as a starting point, we have the following regular alternations:

(32) Regular alternations of the thematic vowel:
   a. S3, S4, S8 have the same thematic vowel as S9.
   b. S2 and S5 have the same thematic vowel as S9 if the TV= a/e.
      If S9 has TV=i, then S2 and S5 have TV=e.
   c. If S9 has TV=a, then S1 has TV=e.
      If S9 has TV=e/i, then S1 has TV=a.
   d. If S9 has TV=a, then S6 has TV=á.
      If S9 has TV=e/i, then S6 has TV=í.
   e. If S9 has TV=a, then S7 has TV=á followed the augment -b-.
      If S9 has TV=e/i, then S7 has TV=í.
   f. If S9 has TV=a, then S10 has TV=á.
      If S9 has TV=e/i, then S10 has TV=í.
   g. If S9 has TV=a, then S11 has TV=a followed by the augment -d-.
      If S9 has TV=e/i, then S11 has TV=i followed by the augment -d-.

Given the regularities between the stems of regular verbs, Spanish verbs only instantiate the following suppletion patterns for the 11 stems of the stem-space. In the following table the related stems are marked by identical letters, cells containing isolated stems are greyed out.

(33)
The predicted possible irregularities are much greater than the irregularities observed. The verb distinguishing the maximal number of stems has 6 stems (ser). If we assume that a verb can maximally have 6 lexically different stems, a theory that does not group together suppletion zones predicts approximately $6.78 \times 10^{43}$ possible suppletions. Assuming a stem-space with 11 suppletion zones (à la Pirrelli & Battista 2000), the possible suppletion patterns with up to 6 lexically different stems are reduced to 601,492, a number which remains much greater than the 12 patterns observed (counting 1 regular and 11 irregular verbs types)\(^{10}\).

Bonami & Boyé (2002, 2003) therefore propose that some organisation of the stem-space further limits the distribution of independent stems inside the stem-space. Bonami & Boyé propose that the stems are connected in an acyclic graph.\(^{11}\) Suppletion spreads only to connected nodes within the graph. Take the following hypothetical examples:

(34) a. graph:

```
A  B  C
```

b. possible suppletion patterns:

```
A  B
A  B
A  B
```

c. impossible suppletion pattern:

```
A  B  A
```

The hypothesis that (34c) is an impossible suppletion pattern corresponds to the intuition that the relations in the graph are strictly local: once a suppletive stem has been lexicalised, the stems attached to the suppleted stem cannot “look through” that stem to recover the regular stem relations with a non-adjacent stem.

The information contained in the graph is suppletion coherence: if two stems are related they cannot be connected through unrelated stems. Two cases are particularly useful to establish connections in the graph:

- First, if all stems but one are related, the isolated stem is at an extremity of the graph, as a result the isolated stem is attached to a single other stem.

---

\(^{10}\) For the mathematical formula used to calculate the number of possibilities see Bonami & Boyé (2003).

\(^{11}\) This is what graph-theory calls a tree. We do not use the term to avoid confusion with trees defined as a structure with two partial orders (dominance, precedence) which are more familiar to linguists.
Secondly, if there are exactly two stems with regularly related values, we can conclude that the two stems are directly linked.

Taking the information from Table (33), this gives us the following information about the structure of the overall graph:

(35)  
   a. S1 (=Present Subjunctive / 1sg Present Indicative) is an isolated stem (caer).
   b. S11 (=Past participle) is an isolated stem (abrir).
   c. S6 (=Preterit / Imperfect 1&2 and Future subjunctive) is an isolated stem (estar).
   d. S2 (=2-3sg Present Indicative/2sg Imperative), and S5 (=3pl Present Indicative), are directly linked in the stem graph.

Furthermore, verbs that have several suppletion zones that are based on the same suppletive stem define sub-graphs that have to be connected:

(36)  
   a. The four stems Future/Conditional, Infinitive/2pl Imperative, Past Participle and Present participle form a sub-graph (ser).
   b. The three stems 1pl Present Indicative, 2pl Present Indicative, 3pl Present Indicative form a sub-graph (ser).

The suppletion patterns in Table (33) do not correspond to a single possible stem-graph. For example consider the sub-graph implied by (36a). The sub-graph can be either one of the graphs in (37):

(37)  
   a. 

   b. 

One possible graph that allows the attested patterns of stem suppletion given in Table (33) is the following:
In the absence of independent criteria that would allow us to choose between competing stem-graphs, we will assume this graph for the sake of argument. Independently of the exact geometry of the stem-graph, the hypothesis of a stem-graph restricting the suppletion patterns of the 11 stems of the Spanish verb stem-space radically reduces the number of predicted suppletion patterns. Assuming maximally 6 distinctly lexicalised stems, any stem-graph that covers the observed relations between stems predicts only 638 possible patterns. This eliminates 99.99% of the 601,492 suppletion patterns predicted by a theory based exclusively on a stem-space with 11 suppletion zones.

As we have said above, this graph captures the regularities of the attested patterns of suppletion; its geometry is therefore motivated by the irregular verbs which by definition have at least two lexically specified stems. The graph in (39) gives the stem-graph of ser: For each of the 6 lexicalised stems, the related stems based on the same lexical stem occupy a connected subgraph:

![Diagram of the stem-graph for ser]

While the irregular verbs give the connected zones of the graph, the regular verbs define the relations that hold between adjacent stems. For regular verbs, exemplified here by amar, the links between stems in the stem-graph always correspond to the regular stem allomorphies given in (32) above:
Given the analysis here, all the stems of a regular verb are related, but not necessarily identical. The question now arises which relations predict the stems of a regular verb given only one lexicalised stem, this means that we have to define a relation for each arrow in the following graph:

(41)

Elaborating on the relations already given in (32), we arrive at the following relations for regular verbs:

(42)  

a. S4 and S9 are identical.  
b. S8 and S9 are identical.  
c. S2 and S5 are identical.  
d. S4 and S3 are identical.  
e. If S5 has TV=a, then S1 has TV=e.  
   If S5 has TV=e/i, then S1 has TV=a.  
f. If S4 has TV=a, then S6 has TV=á.  
   If S4 has TV=e/i, then S6 has TV=í.  
g. If S9 has TV=a, then S7 has TV=á followed the augment -b-.  
   If S9 has TV=e/i, then S7 has TV=í.  
h. If S9 has TV=a, then S10 has TV=á.  
   If S9 has TV=e/i, then S10 has TV=ié.  
i. If S9 has TV=a, then S11 has TV=a followed by the augment -d-.  
   If S9 has TV=e/i, then S11 has TV=i followed by the augment -d-.  
j. If S4 has TV=a, then S5 has TV=a.  
   If S4 has TV=e/i, then S5 has TV=e.
The first four relations are symmetrical. The relations e.-j., however, neutralise the distinction between the thematic vowels -e- and -i-. It is only for the thematic vowel -a- that these relations are clearly reversible. This tallies well with the observation that native speakers of Spanish hesitate between thematic vowels -e- and -i- when confronted with an unknown verb in a form based on a stem that neutralises the distinction between the two:

(43)  (ellos) domben --> (nosotros) dombemos / dombimos.

Notwithstanding, a reverse function could be constructed using the analysis proposed by Boyé & Cabredo Hofherr (2004), which allows to predict the choice between thematic -e- and -i- on the basis of the pre-thematic vowel.

The following graph shows the regular relations between adjacent stems that we propose:12

![Graph showing regular relations between adjacent stems]

We now have a graph that encodes both suppletion coherence for irregular verbs and regular stem allomorphies for regular verbs. Notice that the graph is not directed: for each pair of adjacent stems we define relations in both directions.

Assuming a non-directed graph, irregular verbs are now irregular in two respects:
(i) they lexicalise more than one stem
(ii) they do not obey some of the regularities encoded by the links between stems.

As proposed by Giraud (2005), this can be interpreted as a distinction between active and inactive links: the suppletive stems spread inside the graph only along the active links.

In the graph below, we have marked the active links for haber by full arrows and the inactive links by dashed arrows.

---

12 The relations proposed are noted on the arrows. The identity function is noted by the sign “=”. The functional relations are noted by a “=>”, the non-functional relation by a “-->”.
In the present analysis, in contradiction to Boyé (2000, chapter 6), we have chosen a non-directed graph to account for the intuition that native speakers can construct a novel regular verb based on a single inflected form, independently of the form chosen.

On the one hand, a directed graph would have avoided the additional assumption that in irregular verbs certain links of the graph have to be inactive: in a directed graph a lexicalised stem spreads along the directed links until it reaches a node that has been independently lexicalised.

On the other hand, however, the hypothesis of a directed graph corresponds to a theory in which the reverse relations (even if underspecified in parts) do not form part of the native speaker’s knowledge of Spanish verbal inflection. If confronted with an inflected form such as domber and dombido the native speaker should then be able to generalise the paradigm from one of the forms but not from the other, this does not seem to be the case.13

Conclusion

In the present paper we have proposed an analysis which accounts for the structure of allomorphy of Spanish verbal inflection. This analysis is based on the traditional observation that certain verbal forms always share a common stem. This observation has been generalised by Pirrelli & Battista (2000) to a partition of the verbal paradigm into zones that are systematically coherent under suppletion, giving a stem-space. Following Bonami & Boyé (2003) we have explored the idea that the stems are further organised into a stem-graph that captures the distribution of suppletion within the stem-space. We proposed that the relations within the stem graph are defined by the relations entertained by the stems of the regular verbs. As shown in the graph in (44) these relations are not always the identity relation. Being a regular verb therefore does not reduce to having an invariant stem across the verbal paradigm but rather to obeying the regular pattern of stem allomorphies.

References


13 See Bonami & Boyé (to appear) for a similar discussion for French conjugation.


