Chapter 6

Variable use of strong preterites
A sociolinguistic and theoretical approach

Pilar Barbosa¹, Ana Bastos-Geeii and Cristina Flores¹
¹University of Minho, Portugal / ²University of Connecticut, United States

This paper examines a particular case of variable syncretism in the Minho region in Portugal, involving the 1st person singular (1sg) and 3rd person singular (3sg) forms of “strong” preterits of the verbs *estar* ‘be-stative’, *ter* ‘have’, *fazer* ‘do’ and *ser/ir* ‘be/go’. These forms can be levelled, affecting the 1st or the 3rd person. Fifty interviews from a socially stratified corpus of the relevant dialect were examined by running a mixed effect binominal analysis, which identified as main predictors the factors subject expression, verb and level of education. Moreover, there is a difference between *estar/ter/fazer* ‘be-stative/have/do’ and *ser/ir* ‘be/go’. In the latter case, only one form is used (the 3rd person form *foi*), while in the case of *estar/ter/fazer* ‘be-stative/have/do’ levelling can be realized by either the form for 1sg or 3sg. This inter-linguistic variation is analysed following the “late insertion” model of Distributed Morphology (Halle & Marantz, 1993). We develop an account of these agreement levelling effects that is based on the interaction between the internal syntax of strong preterites and the late insertion of underspecified functional Vocabulary Items. We propose a derivation of the different forms in the standard dialect and then we offer an analysis of levelling where intra-speaker variation is tied to the probabilistic application of feature deleting Impoverishment operations along the lines of Nevins and Parrott (2010). Inter-speaker variation is due to different choices as to which feature sets are subject to Impoverishment: the features for Person or T. For *estar/ter/fazer* ‘be-stative/have/do’ these two operations yield different outputs (in the case of *ter, /teve/ and /tive/, respectively). For *ser/ir* ‘be/go’ the resulting forms are homophonous, namely /foj/ in both cases.

Keywords: verb morphology, strong preterites, paradigm levelling, distributed morphology
1. Introduction

In European Portuguese (EP), in a restricted set of verbs that exhibit stem allomorphy in the preterite (so-called “strong preterites”), the height of the root vowel is the only indication of the category ‘person’ in the forms for 1st person singular (1sg) and 3rd person singular (3sg) (Vivas, 2010). Verbs showing this alternation are estar ‘be-stative’, ter ‘have’, fazer ‘make’, poder ‘can’, pôr ‘put’, and also ir ‘go’ and ser ‘be’, which have identical target forms (see Table 1). In every case, the 1sg is marked by the high vowels [i/u] and the 3sg by the mid-high vowels [e/o].

Table 1. Verbs with vowel alternation in the 1sg and 3sg of the past tense in EP

<table>
<thead>
<tr>
<th>estar</th>
<th>fazer</th>
<th>ter</th>
<th>pôr</th>
<th>poder</th>
<th>ser /ir</th>
</tr>
</thead>
<tbody>
<tr>
<td>estive</td>
<td>fiz</td>
<td>tive</td>
<td>pus</td>
<td>pude</td>
<td>fui</td>
</tr>
<tr>
<td>estiveste</td>
<td>fizeste</td>
<td>tiveste</td>
<td>puseste</td>
<td>pudeste</td>
<td>foste</td>
</tr>
<tr>
<td>estevem</td>
<td>fazem</td>
<td>teve</td>
<td>pos</td>
<td>pode</td>
<td>foi</td>
</tr>
<tr>
<td>estiverem</td>
<td>fizerem</td>
<td>tiveram</td>
<td>puseram</td>
<td>puderam</td>
<td>foram</td>
</tr>
</tbody>
</table>

The present study focuses on a particular case of variation concerning these forms which has been observed in the Portuguese spoken in the Minho region in Portugal. In this region, the 1sg and 3sg forms of strong preterites can be levelled and the form used for levelling can vary: it can be the 1st person (cf. (1)) or the 3rd (cf. (2)).

(1) **1st person instead of 3rd person:**

Pois ela estive lá três anos parada
anyhow she was.1sg there three years still
‘Anyhow she was there still for three years’

(2) **3rd person instead of 1st person**

Eu foi para a tropa em cinquenta e quatro
I went.3sg to the troop in fifty and four
‘I joined the army in fifty four’

Examples (1) and (2) are taken from the corpus that will be analysed in this study. The goals of this paper are: (a) to describe the variable use of the 1sg and 3sg forms of strong preterites in the dialect in question; (b) to determine which linguistic and social factors favor this variation, and (c) to offer an explanatory account of this phenomenon. In Section 2 we present the methodology used, Section 3 shows the results and in Section 4 we explain our theoretical account for the variable patterns observed.
2. Corpus and coding

For the present study, we used the socially stratified corpus of oral discourse Sociolinguistic Profile of the Speech of Braga (cf. the essays in this volume by Aguiar & Paiva, Barbosa, Paiva & Martins, and Rato, Rodrigues, & Varanda). The corpus includes one hour interviews distributed according to four age groups, four education levels and gender. We selected 50 samples from three age groups (26–59 years / 60–75 years / older than age 75) and four different education levels (without a degree / between 4 and 9 years of education / between 10 and 12 years of education / university degree). The variable ‘gender’ was ignored. The 50 speakers of the sub-sample were selected randomly in order to obtain a balanced number of speakers per cell. We aimed to have five speakers per cell, with exception made only to cases in which the main corpus did not have that many speakers for a certain social variable; for instance, it is rare to find speakers with no degree of education among younger people. Table 2 shows the speakers’ distribution according to the selected social variables.

Table 2. Sample: distribution per age span and level of education

<table>
<thead>
<tr>
<th>Level of education</th>
<th>Age span 26–59 years</th>
<th>Age span 60–75 years</th>
<th>Age span older than age 75</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>without degree</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>4–9 class</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>10–12 class</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>university degree</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>18</td>
<td>18</td>
<td>50</td>
</tr>
</tbody>
</table>

The verb forms selected for this study are estive/esteve ‘1sg/3sg of be’, tive/teve ‘1sg/3sg of have’, fiz/fez ‘1sg/3sg of do’, fui/foi ‘1sg/3sg of be/go’, pus/pôs ‘1sg/3sg of put’ and pude/pôde ‘1sg/3sg of can’.

A total of 3,627 tokens were coded, following five variables. ‘Age’ and ‘level of education’ constituted the two social variables. Additionally, three linguistic variables were defined: ‘verb’, ‘subject expression’ and ‘subject position’. The factor ‘verb’ included five different verbs: estar ‘be’, fazer ‘do’, pôr ‘put’, ter ‘have’, poder ‘can’ and ser/ir ‘be/go’ (the latter were coded together). In the condition ‘subject expression’, we considered the values realized subject (Example (3a)), null subject

---

1. The organization of this corpus was funded by FEDER funds through the operational program Factores de Competitividade – COMPETE and by national funds through FCT – Fundação para a Ciência e a Tecnologia, in the context of the Project FCT PTDC/CLE-LIN/112939/2009, Perfil sociolinguistico da fala bracarense. For details, see: <https://sites.google.com/site/projectofalabracarense/>
(Example (3b)) and other subject realizations, which included expletive subjects, interrogative, reflexive and relative pronouns, (Example (3c)). The subject position is either pre – or postverbal (see (4a/b)).

(3) a. \textit{Eu foi para trabalhar de preto, digamos assim…} \[02H1B\]
\begin{flushright}
I went.3sg in.order to.work in black say so
\end{flushright}
‘I went to work wearing black,…’

b. \textit{Fiz a obrigaç\~ao dele.} \[79M4A\]
\begin{flushright}
did.1sg the obligation of.his
\end{flushright}
‘He did his obligation.’

c. \textit{Que há um homem que já estive preso, mata…} \[80M4A\]
\begin{flushright}
That exists a man who already was.1sg arrested kills
\end{flushright}
‘That there is a man who was already in prison, he kills…’

(4) a. \textit{Não sei se ele fiz aquilo, se ele fiz aquilo por vaidade…} \[25H3B\]
\begin{flushright}
Not know if he did.1sg that, if he did.1sg that for vanity
\end{flushright}
‘I don’t know if he did that, if he did that for vanity.’

b. \textit{E então que fez eu?} \[69M3A\]
\begin{flushright}
And then what did.3sg I
\end{flushright}
‘And then what did I do?’

For the statistical analysis we used the variable rule program Rbrul, a text-based interface to existing functions in the R environment (Johnson, 2009). Results are expressed in the level of significance of the predictive variable (p-value), the centered factor weight and log-odds units.$^{2}$

3. Results

In a first step we identified all shifted forms (where the 1st person form was used instead of the 3rd or vice-versa) and all the speakers that presented this variation at least once. The sample contained a total of 195 target forms, i.e., 5.40% of all coded verb forms showed vowel substitution, distributed among 29 speakers. This means that 21 speakers did not show this variation at all.

In a second step we excluded the 21 speakers who never produced shifted forms, reducing the sample to 2232 coded items. Table 3 shows the total occurrences of each verb form and the number of ‘shifted forms’ in each case.

$^{2}$ Log-odds are obtained “from probabilities by taking the natural (base e) logarithm of the odds, where the odds are the probability of an event occurring, divided by the probability of it not occurring” (Johnson, 2009:361).
Table 3. Raw occurrences: total and shifted forms

<table>
<thead>
<tr>
<th></th>
<th>estever</th>
<th>estive</th>
<th>fez</th>
<th>fizer</th>
<th>fizer</th>
<th>fizer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>total</td>
<td>shift</td>
<td>total</td>
<td>shift</td>
<td>total</td>
<td>shift</td>
</tr>
<tr>
<td>68</td>
<td>10</td>
<td>123</td>
<td>16</td>
<td>1289</td>
<td>159</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>1289</td>
<td>119</td>
<td>206</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As can be seen in Table 3, the forms pôde 'could.3sg', pus 'put.1sg' and even pude 'could.1sg' are less frequently used than the other forms. There are only three total occurrences of pôde 'could.3sg' and of pus 'put.1sg'. The form pude 'could.1sg' was never used. The form pôs 'put.3sg' occurs 16 times. Due to the reduced number of tokens, which could bias the results, we decided, in a third step, to exclude these verb forms.

The reduced sample contained a total of 189 shifted forms, which correspond to a mean of 9.40% occurrences, produced by 28 speakers (an additional speaker was excluded who produced twice the excluded form pôde 'could.3sg'). Then, we ran a generalized linear mixed model (GLMM) with the five fixed (independent) factors presented in the previous section. The speakers were defined as a random intercept factor. The dependent variable corresponds to the binary response 'shifted' or 'non-shifted' realization of the past form.

The regression analysis identified as main predictors for this variable use the factor 'speaker', which has a random effect, the linguistic variables subject expression (p < 0.001) and 'verb' (p < 0.001) and the social variable education level (p < 0.001). The variables subject position and age were not selected as predictors. Table 4 presents the statistical results regarding the predictive social variable 'level of education', including log-odd values, number of tokens, proportion of shifted forms and centered factor weight.

Table 4. Statistical output: predictive social variables (level of education)

<table>
<thead>
<tr>
<th>Level of education</th>
<th>Log-dds</th>
<th>Tokens</th>
<th>Proportion of shifted forms</th>
<th>Lefted factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>without degree</td>
<td>1.668</td>
<td>859</td>
<td>12.1%</td>
<td>0.841</td>
</tr>
<tr>
<td>4–9 class</td>
<td>0.375</td>
<td>1013</td>
<td>4.3%</td>
<td>0.593</td>
</tr>
<tr>
<td>10–12 class</td>
<td>−1.535</td>
<td>846</td>
<td>0.7%</td>
<td>0.177</td>
</tr>
<tr>
<td>university degree</td>
<td>−0.508</td>
<td>877</td>
<td>4%</td>
<td>0.376</td>
</tr>
</tbody>
</table>
Regarding the social variables, the results show that variation in this domain tends to occur among the speakers without an education degree. The mean rate of substitution in this group reaches 12.10% and has a centered factor weight of 0.841. ‘Age’ is not a significant predictor, which indicates that this phenomenon is not under progressive change. It appears to represent stable variation that occurs mostly among illiterate speakers. Nevertheless, it has to be mentioned that presently it is rare to find illiterate speakers in younger age spans (younger than 25 years), so there is an overlap between illiterate and older speakers.

Table 5 shows the statistical output concerning the linguistic predictors.

<table>
<thead>
<tr>
<th>Subject expression</th>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
<th>Proportion of shifted forms</th>
<th>Centered factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>realized</td>
<td>0.553</td>
<td>952</td>
<td>11.4%</td>
<td>0.635</td>
<td></td>
</tr>
<tr>
<td>other</td>
<td>0.507</td>
<td>78</td>
<td>9%</td>
<td>0.624</td>
<td></td>
</tr>
<tr>
<td>null</td>
<td>-1.060</td>
<td>2565</td>
<td>2.8%</td>
<td>0.257</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Verb</th>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
<th>Proportion of shifted forms</th>
<th>Centered factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>teve</td>
<td>0.854</td>
<td>157</td>
<td>7.6%</td>
<td>0.701</td>
<td></td>
</tr>
<tr>
<td>esteve</td>
<td>0.593</td>
<td>107</td>
<td>9.3%</td>
<td>0.644</td>
<td></td>
</tr>
<tr>
<td>estive</td>
<td>0.585</td>
<td>239</td>
<td>6.7%</td>
<td>0.642</td>
<td></td>
</tr>
<tr>
<td>foi</td>
<td>0.428</td>
<td>1912</td>
<td>6.2%</td>
<td>0.605</td>
<td></td>
</tr>
<tr>
<td>tive</td>
<td>0.259</td>
<td>346</td>
<td>3.8%</td>
<td>0.564</td>
<td></td>
</tr>
<tr>
<td>fiz</td>
<td>0.108</td>
<td>235</td>
<td>5.5%</td>
<td>0.527</td>
<td></td>
</tr>
<tr>
<td>fez</td>
<td>-0.886</td>
<td>139</td>
<td>2.9%</td>
<td>0.292</td>
<td></td>
</tr>
<tr>
<td>fui</td>
<td>-1.941</td>
<td>460</td>
<td>0.4%</td>
<td>0.126</td>
<td></td>
</tr>
</tbody>
</table>

The results regarding the variable subject expression show that shifted forms are favoured in contexts in which the subject is overtly realized (11.4%; factor weight: 0.635) and rarely occur with null subjects (2.8%; factor weight: 0.257).

As regards the different verb forms, there is a clear difference between the verbs *estar/ter/fazer* ‘be-stative/have/do’, on one hand, and *ser/ir* ‘be/go’, on the other hand. In the former case, the exponent used for levelling may vary – it may be either the 1sg or the 3sg. In the case of *ser/ir* ‘be/go’, by contrast, a single form, namely the 3sg form *foi* ‘went/was’, is almost exclusively used: there are only 2 occurrences of *fui* ‘was.1sg,went.1sg’ used instead of *foi* ‘was.3sg, went.3sg’ (2/469; 0.40%), against 119 occurrences of *foi* ‘was.3sg, went.3sg’ used instead of *fui* ‘was.1sg,went.1sg’ (119/1912; 6.20%).
Finally we looked at individual variation. Among the 28 speakers who produce shifted forms, 27 speakers show variation with *ser/ir 'be/go*. Eleven out of these 27 speakers only show variation with this verb. Seventeen speakers show variation with two or more verbs. Additionally, there is a marked tendency of the speakers to use a shifted form only in one way per verb. To put it differently, a given speaker either uses the 3rd person instead of the 1st person or vice-versa, but rarely both forms of one verb randomly. There are only two speakers that use both forms of a given verb. This is shown in Table 6.

Table 6. Distribution of shifted forms per speaker

<table>
<thead>
<tr>
<th>Speaker</th>
<th>esteve</th>
<th>estive</th>
<th>fez</th>
<th>fiz</th>
<th>foi</th>
<th>fui</th>
<th>teve</th>
<th>tive</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fal12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Fal13</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Fal14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Fal16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Fal18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Fal22</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Fal25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Fal31</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Fal33</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Fal34</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Fal35</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Fal38</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Fal44</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Fal55</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Fal59</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Fal62</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Fal67</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>Fal68</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>2</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>Fal69</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Fal70</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Fal72</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Fal74</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Fal79</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>Fal80</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Fal81</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Fal82</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Fal84</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Fal90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>16</td>
<td>4</td>
<td>13</td>
<td>119</td>
<td>2</td>
<td>12</td>
<td>13</td>
<td>189</td>
</tr>
</tbody>
</table>
The consistent use of a given form within the speech of the same speaker leads to the conclusion that variation is not random. In particular, there is inter-individual variation in the choice of the form used for paradigm levelling. Since each individual speaker alternates between the use of the standard form and syncretism, there are two different kinds of variation here: intra-and inter-individual variation.

4. Discussion and theoretical consequences

In this section, we present an analysis of the patterns described above. In view of the results of the preceding section, our goals are twofold: (i) to capture not only inter but also intra-individual variation; (ii) to offer an explanatory account of why there is an asymmetry between fui ‘was.1sg, went.1sg’/ foi ‘was.3sg, went.3sg’ and the other verb forms.

Before we go into the analysis, it is important to notice that we are not dealing with a phenomenon that is purely and exclusively phonetic in nature. Several considerations conspire against this possibility. In the first place, purely phonetic phenomena are indifferent to grammatical class and apply indiscriminately across categories. We do not observe this kind of alteration in vowel quality anywhere else besides this limited class of verbs. This indicates that the phenomenon in question belongs with morphology, or at least to the domain of the morphophonology. In addition, we observed that there is a particular pattern per speaker: each speaker has a particular choice of the form used for levelling with a particular verb. This is another feature that renders a phonetic treatment rather unlikely. Finally, variation is conditioned by the overtness of the subject, an obvious morphosyntactic factor, that certainly doesn’t bear on vowel quality. For these reasons, we will explore an analysis that relies on the idea that agreement levelling is due to operations that take place at a level of the grammar that stands at the interface between syntax and morphology.

So as to attain the goals stated above, we will explore an idea originally put forward in Nevins and Parrott (2010), according to which intra-individual paradigm levelling is best analysed as the probabilistic application of post-syntactic feature-deleting ‘Impoverishment’ operations (Bonet, 1991). Such operations have been well established within the model of Distributed Morphology (DM) (Bonet, 1991; Halle, 1997; Harley, 2008; Noyer, 1998). The novelty of the proposal found in Nevins and Parrott (2010) lies in their use of the notion of “variable rule” taken from Variationist Sociolinguistics (Labov, 1972; Guy, 1991) as a specific mechanism within the theoretical framework of DM.

In what follows, we will attempt to formalize the underlying processes responsible for the patterns observed, using the tools of DM. This section is organized as
Chapter 6. Variable use of strong preterites

follows. In Section 4.1 we briefly present the theoretical framework and our background assumptions; in Section 4.2 we propose a specific analysis of the paradigms for strong preterites in standard EP; the patterns of variation observed in the dialect under analysis are discussed and analysed in Section 4.3.

4.1 Background assumptions

DM assumes the T-model of the architecture of the grammar, as proposed in the Principles and Parameters framework of Generative Grammar (Chomsky, 1986, 1995 and subsequent work).

(5) Syntactic Derivation
   Spell Out
   Morphological Structure
   Phonetic Form
   Logical Form

DM posits a Morphological Structure (MS) level between Spell Out and PF. In this model, the output of narrow syntax is the input to MS, in which further operations apply during the computation to Phonetic Form. DM is a “late insertion” model. The terminal nodes of the syntactic derivation – labelled ‘morphemes’ – are bundles of abstract features relevant only to syntax and the phonological exponents for functional morphemes are inserted post-syntactically, via the process of Vocabulary Insertion (VI). The ‘vocabulary’ is a list of the phonological exponents of the different abstract morphemes of the language, paired with conditions of insertion (see below).

Following Marantz (1997), we take roots not to have any category; in the syntax they are merged with category-giving functional heads. In the verbal domain, this head is v and is responsible for the verbal properties of the verbal complex, like (in) transitivity, agentivity, (accusative) case, and so on. Along the lines of Oltra-Massuet (1999) for Catalan, we assume that, in the unmarked case, the Tense (T), Mood (M), and Aspect (Asp) features are combined in EP into a single morpheme for the computational system, which we will label T for short. In the syntax, a root merges with the verbalizing head v and further undergoes cyclic head-to-head movement all the way up to T. T also dominates a head hosting agreement features. This results in the following minimal representation for the verbal complex under T:
One further, fairly uncontroversial, assumption we make is that at MS $v$ has a theme vowel ($Th$) adjoined to it.

$Th$ defines the class $v$ belongs to. In EP, there are three verb class features ($Th_1$, $Th_2$, $Th_3$) corresponding to the theme vowels $-a$, $-e$, $-i$, respectively.\textsuperscript{3}

As for the $\phi$-feature set to be inserted under Agr, we adopt Nevins’s (2007) proposal according to which the feature matrices for person and number combine the following features:

(8) $\phi$-features for Person: [$\pm$participant, $\pm$author]  
\textbf{Number features:} [$\pm$plural]

The different combinations of $\phi$-features give us the following specifications for the Agr node:

(9) 1sg: [+part; +auth; −pl]  
2sg: [+part; −auth; −pl]  
3sg: [−part; −pl]  
1pl: [+part; +auth; +pl]  
2pl: [+part; −auth; +pl]  
3pl: [−part; +pl]

\textsuperscript{3} Oltra-Massuet (1999) argues, on the basis of Catalan, that every functional head requires a (possibly null) theme vowel. However, since this issue is not crucial for the EP cases under discussion, we will not include the theme vowel under $T$ in our representations.
As mentioned above, in the PF component of the grammar, the functional morphemes receive phonological representations via VI where each Vocabulary Item is defined as a pair that specifies the relation between a phonological expression and a grammatical or semantic feature and possibly a context of insertion. We propose the following Vocabulary Items for insertion under \(Th\) (9), under \(T\) in the Past and Future tenses, and under \(Agr\) (12). The exponent for the categorizing head \(v\) is invariably zero, so we do not include it in our lists. The Vocabulary Items in (10–12) are adapted from Teixeira (2012) as well as Oltra-Massuet and Arregi (2005).

\[
\begin{align*}
(10) & \quad \text{Thematic vowel} \\
& /a/ \leftrightarrow [Th_1] \\
& /e/ \leftrightarrow [Th_2] \\
& /i/ \leftrightarrow [Th_3] \\
(11) & \quad T \text{ (non exhaustive)} \\
& /va/ \leftrightarrow [−\text{perfect}, +\text{past}] / [Th_1] \\
& /a/ \leftrightarrow [−\text{perfect}, +\text{past}] / [Th_2] \text{ or } [Th_3] \\
& /r/ \leftrightarrow [+\text{fut}] \\
(12) & \quad \varphi\text{-features} \\
& p>/i/ \leftrightarrow [+\text{perf}, +\text{past}, +\text{part}, +\text{auth}, −\text{pl}] \\
& /s/ \leftrightarrow [+\text{part}, −\text{auth}, −\text{pl}] \\
& /mos/ \leftrightarrow [+\text{auth}, +\text{pl}] \\
& /N/ \leftrightarrow [+\text{pl}] \\
& /Ø/ \leftrightarrow [−\text{pl}] \\
& /u/ \leftrightarrow [+\text{perf}, +\text{past}, −\text{pl}] \\
& /ste/ \leftrightarrow [+\text{perf}, +\text{past}, +\text{part}, −\text{auth}, −\text{pl}] \\
& /o/ \leftrightarrow [+\text{pres}, +\text{auth}, −\text{pl}] \\
& /ra N/ \leftrightarrow [+\text{perf}, +\text{past}, +\text{pl}] \\
\end{align*}
\]

One important feature of late insertion is under specification, which is governed by the Subset Principle (Halle, 1997). This principle determines that Vocabulary Items do not need to be fully specified for insertion at a particular node:

\[
(13) \quad \text{Subset Principle} \\
\text{The phonological exponent of a vocabulary item is inserted into a position if the item matches all or a subset of the features specified in that position. Insertion does not take place if the Vocabulary Item contains features not present in the morpheme.}
\]

---

4. The items listed are those proposed by Teixeira (2012) with one difference, namely the last one listed under (11): \(/raN/ \leftrightarrow [+\text{perf}, +\text{past}, +\text{pl}]\). The reasons for including this item will become clear below (see Footnote 5).
The following Principle regulates competition among different Vocabulary Items capable of being inserted at a particular node:

(14) **Blocking Principle**
Whenever there is more than one Vocabulary Item capable of being inserted, the more specified item is to be chosen.

We can see how these two principles work when we consider the derivation of the forms of the Imperfective Past of a regular verb such as *cantar* ‘to sing’.

(15) **Imperfective Past**

<table>
<thead>
<tr>
<th>Person</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>cantava</td>
</tr>
<tr>
<td>2sg</td>
<td>cantavas</td>
</tr>
<tr>
<td>3sg</td>
<td>cantava</td>
</tr>
<tr>
<td>1pl</td>
<td>cantávamos</td>
</tr>
<tr>
<td>3pl</td>
<td>cantavam</td>
</tr>
</tbody>
</table>

Starting with the root /kanta/, the theme vowel /a/ is selected for insertion in conjunction to ν and the exponent /va/ is inserted under the node T. The morphemes selected for insertion under Agr are as shown in (16): the Subset Principle and the Blocking Principle together force insertion of /s/ under a node that is specified for 2sg and /mos/ under a node that is specified for 1pl.

(16) a. /ø/ ↔ [−pl] 1sg: [Agr+part; +auth; −pl]
b. /s/ ↔ [+part, −auth, −pl] 2sg: [Agr+part; −auth; −pl]
c. /ø/ ↔ [−pl] 3sg: [Agr−part; −pl]
d. /mos/ ↔ [+auth,] 1pl: [Agr+part; +auth; +pl]
e. /N/ ↔ [+pl] 3pl: [Agr−part; −auth; +pl]

We thus obtain the following outputs:

(17) 1sg  [[[√kanta] -a | T-va ]φ-ø]
2sg  [[[√kanta] -a | T-va ]φ-s]
3sg  [[[√kanta] -a | T-va ]φ-ø]
1pl  [[[√kanta] -a | T-va ]φ-mos]
3pl  [[[√kanta] -a | T-va ]φ-N]

One other process that will play an important role in our analysis is the mechanism of Fusion. This operation merges two terminal nodes under a single node so that only one Vocabulary Item can be inserted. According to Teixeira (2012) (as well as Bassania & Lunguinho, 2011), this process applies in the derivation of the Past Perfect (= Preterite Perfect). The paradigm for the Past Perfect of *cantar* ‘to sing’ is the following:
Chapter 6. Variable use of strong preterites 165

(18) 1SG  
cantei
2SG  
cantaste
3SG  
cantou
1PL  
cantámos
3PL  
cantaram

Teixeira (2012) proposes that the nodes t and agr fuse in the Past Perfect in all persons except 3pl, as illustrated in (19).

(19)

Since a full discussion of the motivation for excluding 3pl from this process is well beyond the scope of the present paper, we will assume that Fusion applies in all the persons of the paradigm. 5

Considering the Vocabulary Items listed in (12), the exponents for insertion under the fused terminal node \( \text{t/agr} \) are as indicated in (20).

(20)  

Terminal node for \( \text{t/agr} \)  
Exponent  
Output

| 1SG | a. \([\text{t/agr}+\text{PERF},+\text{PAST},+\text{PART},+\text{AUTH},−\text{PL}]\) | ↔ /i/ | /ka nt+a +i/ |
| 2SG | b. \([\text{t/agr}+\text{PERF},+\text{PAST},+\text{PART},−\text{AUTH},−\text{PL}]\) | ↔ /ste/ | /ka nt+a +ste/ |
| 3SG | c. \([\text{t/agr}+\text{PERF},+\text{PAST},−\text{PART},−\text{AUTH},−\text{PL}]\) | ↔ /u/ | /ka nt+a +u/ |
| 1PL | d. \([\text{t/agr}+\text{PERF},+\text{PAST},+\text{PART},+\text{AUTH},+\text{PL}]\) | ↔ /mos/ | /ka nt+a +mos/ |
| 3PL | e. \([\text{t/agr}+\text{PERF},+\text{PAST},+\text{PL}]\) | ↔ /ra N/ | /ka nt+a +ra N/ |

The outcomes for the forms for 1SG and 3SG are subject to phonological changes, so that /kant+a+i/ becomes /kantej/ and /kant+a+u/ surfaces as /kantou/. In the DM model, this kind of changes results from the operation of readjustment rules that have the form of phonological rules and apply to morphemes after VI. Teixeira (2012) proposes the following two readjustment rules, which yield the desired outcomes:

---

5. This is why we need to add the Vocabulary Item mentioned in the previous footnote.
(21) **Phonological readjustment rules**

a. /a/ → /e/ _/i(S)/

b. /a/ → /o/ /_/u/

Having established our background assumptions, we now turn to an analysis of Strong Preterites in “standard” EP.

### 4.2 Strong preterites in standard EP

As has already been observed, all of the preterites under discussion exhibit stem allomorphy (see Table 1 above). Within this set of verbs, a major split can be drawn between *estar/ter/fazer* ‘be-stative/have/do’, on the one hand, and *ser/ir* ‘be/go’, on the other. While *ser/ir* ‘be/go’ exhibit truly suppletive forms, *estar/ter/fazer* ‘be-stative/have/do’ do not. We illustrate this point by comparing the forms for the 1sg in different tenses:

(22)

a. *vou* ‘go.pres.1sg’ – *fui* ‘go.perf.past.1sg’ – *ia* ‘go.imperf.past.1sg’

b. *sou* ‘be.pres.1sg’ – *fui* ‘be perf.past.1sg’ – *era* ‘be imperf.past.1sg’

(23)

a. *estou* ‘go.pres.1sg’ – *estive* ‘go.perf.past.1sg’ – *estava* ‘go.imperf.past.1sg’

b. *tenho* ‘be.pres.1sg’ – *tive* ‘be perf.past.1sg’ – *tinha* ‘be imperf.past.1sg’

c. *faço* ‘be.pres.1sg’ – *fiz* ‘be perf.past.1sg’ – *fazia* ‘be imperf.past.1sg’

Unlike the forms in (22), the stems in (23) maintain some integrity and have at least some segments in common. Another (related) feature that distinguishes the two sets of verbs is that *ser/ir* ‘be/go’ are athematic in all forms of the paradigm for the Past Perfect, whereas *estar/ter/fazer* ‘be-stative/have/do’ do contain a thematic vowel, namely /e/, in some persons of the paradigm: 2sg, 1pl and 3pl; the only cases in which the thematic vowel is missing are precisely 1sg and 3sg.

In order to capture the athematic nature of *ser/ir* ‘be/go’ in every person in the Past Perfect paradigm, and of *estar/ter/fazer* ‘be-stative/have/do’ in 1sg and 3sg, we propose that, in these cases, Fusion applies to \( \sqrt{v} \), thus blocking insertion of the thematic vowel. Since \( t \) and \( \text{AGR} \) fuse in the Past Perfect, we have the following configuration for these cases:
In all other persons of the paradigm of *estar/ter/fazer* ‘be-stative/have/do’, only T and Agr fuse; therefore, the thematic vowel /e/ is inserted (the representation is the same as that of a regular verb (cf. (19)).

In addition, we propose the following rules of allomorphy:

(26) **Rules of allomorphy:**

- a. √*ser/ir* ↔ /fo/ ___ [+PERF; +PAST]
- b. √*fazer* ↔ /fis/ ___ [+PERF; +PAST]
- c. √*estar* ↔ /estiv/ ___ [+PERF; +PAST]
- d. √*ter* ↔ /tiv/ ___ [+PERF; +PAST]

With this much as background, let us consider the derivations of the forms for 2SG/1PL/3PL of both sets of verbs. In the case of *estar/ter/fazer* ‘be-stative/have/do’, the relevant representation is the one in (26), so we have the following terminal nodes:

(27) 2SG: \[T/AGR[V \{v[\sqrt{\cdot}] \circ \} Th_2] \; +\text{PERF}; +\text{PAST}; +\text{PART}; −\text{AUTH}; −\text{PL}\]

1PL: \[T/AGR[V \{v[\sqrt{\cdot}] \circ \} Th_2] \; +\text{PERF}; +\text{PAST}; +\text{PART}; +\text{AUTH}; +\text{PL}\]

3PL: \[T/AGR[V \{v[\sqrt{\cdot}] \circ \} Th_2] \; +\text{PERF}; +\text{PAST}; −\text{PART}; −\text{AUTH}; +\text{PL}\]
The root combines with the thematic vowel /e/ and the Vocabulary Items for insertion under $\text{T/AGR}$ are /ste/ for 2sgg, /mos/ for 1pl and /raN/ for 3pl (cf. the list in (12)). This yields the following forms:

\[(28) \quad \begin{array}{lll}
\text{2sg} & \text{1pl} & \text{3pl} \\
\text{Estar} & /\text{estiv}+\text{e}+\text{ste}/ & /\text{estiv}+\text{e}-\text{mos}/ & /\text{estiv}+\text{e}+\text{raN}/ \\
\text{Fazer} & /\text{fis}+\text{e}+\text{ste}/ & /\text{fis}+\text{e}-\text{mos}/ & /\text{fis}+\text{e}+\text{raN}/ \\
\text{Ter} & /\text{estiv}+\text{e}+\text{ste}/ & /\text{tiv}+\text{e}-\text{mos}/ & /\text{tiv}+\text{e}+\text{raN}/
\end{array}\]

In the case of ser/ir ‘be/go’, the relevant representation is the one on the right hand side of (24), so thematic vowel insertion is blocked. Combining the root /fo/ with the Vocabulary Items compatible with the terminal nodes for 2sg (/ste/), 1pl (/mos/) and 3pl (/raN/), the outcome is, respectively, /fo+ste/, /fo+mos/, /fo+raN/.

Now we move on to 1sg and 3sg. Here, the syntactic representation is (24) with both sets of verbs. Starting with 1sg, the terminal node $\text{T/AGR}$ has the feature specification $[+\text{perf}; +\text{past}; +\text{part}; +\text{auth}; -\text{pl}]$. This node is only compatible with the Vocabulary Item that inserts the exponent /i/ (cf. (12)). We get the following underlying forms for each verb:

\[(29) \quad \begin{array}{llll}
\text{a.} & /\text{fo}+\text{i}/ \\
\text{b.} & /\text{tiv}+\text{i}/ \\
\text{c.} & /\text{fis}+\text{i}/ \\
\text{d.} & /\text{estiv}+\text{i}/
\end{array}\]

So as to derive the surface forms, we posit the following Phonological Readjustment rules. The first one was already proposed by Mateus (1982) and applies quite generally in Ep. It centralizes the word final non-stressed high vowel when following a consonant:

\[(30) \quad \text{Centralization (Mateus, 1982)} \\
/\text{i}/ \rightarrow /\text{i}/ \quad \text{C,} \quad #\]

Upon application of this rule, the surface forms obtained for (30) are /estivi/, /tivi/, /fizi/. In the latter case, a further phonological rule deletes the final vowel, yielding /fis/:

\[(31) \quad /\text{i}/ \rightarrow /\emptyset / \quad \text{s,} \quad #\]

The other phonological readjustment rule raises the midvowel of the root /fo/ to /u/ under adjacency with the high vowel /i/:

\[(32) \quad \text{Phonological Readjustment rule} \\
/\text{fo}/ \rightarrow /\text{fu}/ \quad [+\text{perf, +past} - /i/ ]\]
Upon application of rule (32) followed by syllabification and glide formation, the outcome is /fuj/.

Now we move on to 3sg. In the case of regular verbs, the item for insertion under the node [+PERF, +PAST, −PART, −AUTH, −PL] is /u/ (cf. (12)). In the case of the verbs under discussion, what happens instead is that the stressed vowel of the root is realized as a mid-vowel. In order to account for this, we propose a rule of allomorphy that inserts a zero morpheme under a node that is specified as [+PERF, +PAST, −PL] under adjacency with the root of the verbs in question:

(33) \[ T/AGR +PERF, +PAST, −PL \leftrightarrow ø / \[ √/v \{ √estar, √ter, √fazer, √ir, √ser, …\} \]

Since the Vocabulary Item in (33) is more specified than the Vocabulary Item that inserts /u/, it wins the competition. As a result of this, the underlying forms are the following:

(34)  a. [[fo] ø]
    b. [[estiv] ø]
    c. [[tiv] ø]
    d. [[fiz] ø]

For the case of /fo/, we propose that a phonological readjustment rule inserts an epenthetic /e/ yielding the surface form /foe/. Since e-epenthesis is a more general process, we assume it applies whenever necessary.

(35)  e-epenthesis (whenever necessary): ø ↔ /e/

By hypothesis, e-epenthesis is required in the case of /fo/ to create a foot. After syllabification, glide formation applies and the output form is /foj/.

Turning now to the other forms, we propose a further morphophonological readjustment rule, which lowers the stressed vowel of the root whenever followed by T/Agr [+PERF, +PAST − ø]. We formulate the rule in question as follows, after Halle & Marantz (1993):

(36)  Vowel Lowering:
    \[ V \rightarrow [−high] / W ___ U T/Agr[+PERF, +PAST − ø] \# \]
    where wvu = /estiv/, /tiv/, /fis/

As a result of this rule, the forms obtained are /tev/, /estev/, /fes/. Since /s/ is a possible coda in EP, no further changes are required for /fes/. /v/, on the other hand, is not an admissible coda in EP. For this reason, e-epenthesis must apply in the case of /tev/, /estev/, yielding the forms /esteve/, /teve/.
4.3 Variable syncretism

In this section we propose an analysis of the different patterns of paradigm levelling in the dialect under discussion. Our analysis aims to capture not only intra-individual variation, that is, the fact that a given speaker alternates between the target form and a levelled form, but also inter-individual variation: the observation that, in the case of estar/ter/fazer ‘be-stative/have/do’, different forms may be chosen for levelling: either 1sg or 3sg. Furthermore, our analysis should also able to provide an account for why this latter type of variation is unattested in the case of the verbs ser/ir ‘be/go’.

In a recent paper, Nevins and Parrott (2010) revive the notion of ‘variable rule’ from Variationist Sociolinguistics (Labov, 1972; Guy, 1991) as a specific mechanism within the theoretical framework of DM and propose that intra-individual paradigm levelling variation is due to the probabilistic application of feature deleting Impoverishment operations (Bonet, 1991). Impoverishment rules apply in Morphological Structure to morphosyntactic features (that is, prior to Vocabulary Insertion) and eliminate certain distinctions. As a result, the Vocabulary Item that is expected to be chosen can no longer be inserted and a less specified item, generally the elsewhere case, is chosen instead; in other words, there is “retreat to the elsewhere case”. One further claim made by Nevins and Parrott (2010) (see also Oltra-Massuet, 2013) is that morphological Impoverishment operations are induced by the inherent and universal markedness of particular morphosyntactic features or their combination (Croft, 2003; Greenberg, 1966). In particular, they claim that for all three φ-features introduced in Section 4.1 (±Part; ±Auth; ±pl), as well as for Tense (±Past), the positive value + is marked and hence a target for Impoverishment (Greenberg, 1966). Here we adopt these ideas.

Considering that, with respect to the subset of verbs estar/ter/fazer ‘be-stative/have/do’, there is inter-individual variation with respect to the form chosen for levelling (1sg or 3sg), we propose that there are two rules of Impoverishment which are independent from each other.

The first Impoverishment operation targets the φ-features for 1st person, that is, [+part, +author], in the singular:

(37) **Option 1**: Impoverishment of the φ-features for 1st person


\[ [+\text{Auth}, +\text{Part}] \rightarrow \emptyset / \{\sqrt{\text{ser}}, \sqrt{\text{estar}}, \ldots \} |_{T/Agr} ^{-\text{pl}} \]

Rule (37) applies probabilistically. It deletes the positively specified Participant and Author features from the feature matrices specified as [−pl] in the context of √ser, √ir, √estar, √fazer, √ter. The input and the output of the rule are as follows:
(38) a. Input
\[
\begin{array}{c}
T \\
\sqrt{v} \\
T/\text{AGR} \\
[+ \text{Perf}, + \text{Past}, + \text{auth}, + \text{part}, -\text{PL}] \\
\end{array}
\]
b. Output
\[
\begin{array}{c}
T \\
\sqrt{v} \\
T/\text{AGR} \\
[+ \text{Perf}, + \text{Past}, -\text{PL}] \\
\end{array}
\]

As we have seen, allomorphy determines that the exponents for $\sqrt{v}$ are /fo/, estiv/, /tiv/, /fis/. As for the exponents for insertion under $T/\text{AGR}$, we have the following items in competition:

(39) a. $\varnothing \leftrightarrow [+\text{perf}, +\text{past}, -\text{pl}] / \{\sqrt{estar}, \sqrt{ter}, \sqrt{ir}, \sqrt{ser}\}$
b. /u/ $\leftrightarrow [+\text{perf}, +\text{past}, -\text{pl}]$
c. /i/ $\leftrightarrow [+\text{perf}, +\text{past}, +\text{part}, +\text{auth}, -\text{pl}]$
d. $\varnothing \leftrightarrow [-\text{pl}]$

Since the most richly specified item is (39a), we have the following forms:

(40) a. ser/ir: [[fo] $\varnothing$]
b. ter: [[tiv] $\varnothing$]
c. estar: [[estiv] $\varnothing$]
d. fazer: [[fis] $\varnothing$]

The surface forms are derived as follows:

(41) **Ser/ir**
   Input: [[fo] $\varnothing$ ]
   e -epenthesis: /foe/
   Syllabification and glide formation: /foj/
   Output: /foj/

(42) **Estar/ter/fazer**
   Input: [[tiv] $\varnothing$]; [[estiv] $\varnothing$]; [[fis] $\varnothing$]
   Vowel lowering (rule (36)): /tev/; /estev/; /fes/.
   Syllabification and epenthetic vowel insertion: /teve/; /esteve/
   Ouput: /teve/; /esteve/; /fes/
These output forms are homophonous with the form for 3sg, so we derive *eu foi*, *eu teve*, *eu esteve* and *eu fez*.

Now we turn to the derivation of the opposite pattern: the use of 1sg instead of 3sg. Keeping with the insight that Impoverishment rules target positively specified features, we propose that it is the tense features that are subject to deletion, as follows:

(43) **Option 2: T Impoverishment**

\[ [+\text{Pret Perf}] \rightarrow \emptyset / \{\sqrt{\text{ser}}, \sqrt{\text{estar}}, \ldots\}[_{T/\text{AGR}} \text{ -- PL}] \]

T impoverishment is illustrated in (44)

(44) a. **Input**

```
T
/\v
T/AGR
[+ Perf, + Past, + auth, + part, – PL]
```

b. **Output**

```
T
/\v
T/AGR
[+ auth, + part, – PL]
```

The exponents for insertion under \( \sqrt{v} \) are /fo/, /estiv/, /tiv/, /fis/. The items that compete for insertion under T/AGR are listed in (45); item (45d) is the winner.

(45) **Items in competition:**

a. /u/ \( \leftrightarrow \) [+PERF, +PAST, – PL]
b. /o/ \( \leftrightarrow \) [+PRES, +PART, +AUTH, – PL]
c. /i/ \( \leftrightarrow \) [+PERF, +PAST, +PART, +AUTH, – PL]
d. \( \emptyset \) \( \leftrightarrow \) [– PL]
e. \( \emptyset \) \( \leftrightarrow \) [+PERF, +PAST, – PL] / {\sqrt{estar}, \sqrt{ter}, \sqrt{ir}, \sqrt{ser}}

After Vocabulary Insertion, we get the following structures:

(46) a. *ter*: [[/tiv/] \( \emptyset \)]
b. *estar*: [[estiv] \( \emptyset \)]
c. *fazer*: [[fis] \( \emptyset \)]
d. *ser/ir*: [[fo] \( \emptyset \)]
The final output is derived as follows:

(41)  
Ser/ir  
Input: [[fo] ø ]  
e -epenthesis: /foe/  
Syllabification and glide formation: /foj/  
Output: /foj/  

(42)  
Estar/ter/fazer  
Input: [[tiv] ø]; [[estiv] ø]; [[fis] ø]  
Syllabification: /tiv/; /estiv/; /fis/  
Epenthetic vowel insertion: /tive/; /estive/  
Output: /tive/; /estive/; /fis/  

The output forms and the forms for 1sg are syncretic in the case of estar/ter/fazer ‘be-stative/have/do’, but not in the case of ser/ir ‘be/go’, which is the result we want. Thus, T-impoverishment and φ-impoverishment yield different outcomes only in the case of the set of verbs estar/ter/fazer ‘be-stative/have/do’. In these cases, T-impoverishment yields output forms that are identical to the forms for 1sg; φ-impoverishment results in strings that are identical to the forms for 3sg. As regards ser/ir ‘be/go’, the output form is /foj/ in either case.

We have seen that there is a strong tendency for a given speaker to follow only one pattern. This is accounted for under the assumption that only one of the rules is part of the internalized grammar of a given speaker for a given verb. The existence of the two options, however, accounts for inter-individual variation. The hypothesis that both rules of Impoverishment apply in a probabilistic manner accounts for intra-individual variation, that is, the fact that speakers fluctuate between the forms of the standard grammar and syncretism.

5. Final remarks

In this paper, we have examined two particular patterns of variable paradigm levelling in strong preterites in European Portuguese. A quantitative analysis of the occurrence of this kind of syncretism in a socially stratified speech corpus shows that variation is conditioned by (i) the overtness of the subject; (ii) level of education. In view of (i), we have concluded that the phenomenon in question stands at the interface between the syntax and morphology.

In our analysis, we followed the “late insertion” model of Distributed Morphology (Embick & Noyer, 2007; Halle & Marantz, 1993), which assumes the T-model of the architecture of the grammar as proposed in the Principles and
Parameters framework of Generative Grammar (Chomsky, 1986, 1995 and subsequent work). Adopting standard syntactic structures and following basic assumptions about the structure of the Portuguese verb (Oltra-Massuet & Arregi, 2005; Teixeira, 2012), we developed a Distributed Morphology account of the agreement levelling effects noted that is mainly based on the interaction between the internal syntax of the strong preterites and the late insertion of underspecified Vocabulary items. We proposed a derivation of the different forms in the standard dialect and then we offered an analysis of levelling where intra-speaker variation is tied to the probabilistic application of Impoverishment rules along the lines of Nevins & Parrott (2010) and Oltra-Massuet (2013). Inter-speaker variation is due to different choices as to which feature sets are subject to Impoverishment: [+Author; +Participant] (= 1st person) or T. In the case a verb such as ter, this yields /teve/ and /tive/, respectively. In the case of ser or ir the resulting form is /foj/ in both cases.

References


