

Faculdade de Letras da Universidade de Coimbra



**DESVIOS NA REPRESENTAÇÃO ORTOGRÁFICA DAS  
VOGAIS POSTERIORES POR APRENDENTES DE  
PORTUGUÊS LÍNGUA NÃO MATERNA (PLNM)**

Matilde Bernardes Pinto

Coimbra - 2011

Faculdade de Letras da Universidade de Coimbra



**DESVIOS NA REPRESENTAÇÃO ORTOGRÁFICA DAS  
VOGAIS POSTERIORES POR APRENDENTES DE  
PORTUGUÊS LÍNGUA NÃO MATERNA (PLNM)**

Matilde Bernardes Pinto

Dissertação de Mestrado em Português Língua Estrangeira/ Língua Segunda  
apresentada à Faculdade de Letras da Universidade de Coimbra

Orientadoras: Professora Doutora Cristina Santos Pereira Martins  
Professora Doutora Isabel Maria Almeida Santos

Coimbra - 2011

# Agradecimentos

Agradeço às Professoras Doutoradas Cristina Santos Pereira Martins e Isabel Maria Almeida Santos, da Faculdade de Letras da Universidade de Coimbra, por terem orientado esta dissertação e se terem mostrado, ao longo de todo o trabalho, sempre disponíveis.

Agradeço igualmente aos meus pais, irmão, namorado e amigos pelo apoio e incentivo que sempre me deram ao longo deste percurso, sem nunca me deixarem desanimar, pela paciência demonstrada nos momentos mais angustiantes, pelo interesse que sempre demonstraram e pela ajuda prestada, essencial para alcançar mais esta meta.

Agradeço em particular aos meus pais por todo o apoio, emocional e material, necessário ao longo deste tempo. Sem eles não teria sido possível.

A todos, um enorme obrigada.

## Resumo

Este trabalho pretende analisar os erros envolvendo vogal posterior nas produções escritas por parte de aprendentes de português língua não materna. Estes aprendentes frequentavam o *Curso Anual de Língua e Cultura Portuguesas para Estrangeiros*, o *Curso de Férias de Língua e Cultura Portuguesas para Estrangeiros* e o curso de *Língua Portuguesa Erasmus* leccionados na Faculdade de Letras da Universidade de Coimbra. Para este estudo analisou-se, então, o *corpus* de Produções Escritas de Aprendentes de Português L2 (PEAPL2/CELGA), composto pelos textos produzidos pelos referidos aprendentes. Assim, tendo sido estudados todos os textos do *corpus*, seleccionaram-se aqueles que tinham erros ortográficos envolvendo vogal posterior e procedeu-se à sua análise.

A primeira parte da dissertação é constituída por dois capítulos de enquadramento do tema, nos quais se procede à descrição do sistema vocálico do português europeu (PE) na sua variedade padrão e se tecem considerações sobre as relações entre as unidades posteriores e as suas representações ortográficas.

Na segunda parte são analisados os erros em causa de acordo com a tipologia estabelecida para o efeito, levando em consideração o nível de aprendizagem de PL2 dos informantes.

Por fim, discutem-se as causas que poderão ter levado os aprendentes ao erro, na tentativa de fornecer informações relevantes para o desenvolvimento de estratégias que ajudarão a uma redução destes desvios.

## Abstract

This thesis is focused on analyzing the errors that are involved in the written representation of back vowels by those who learn Portuguese as a foreign language (PL2). The subjects in this study were all students of either the Annual Course of Portuguese Language and Culture for Foreigners, the Summer Course of Portuguese Language and Culture for Foreigners or the ERASMUS Portuguese Course taught in the Faculty of Arts, in the University of Coimbra. For this study, written texts produced by learners of Portuguese as L2 (*Corpus PEAPL2/CELGA*) were examined. From these texts, all those exhibiting misspellings of back vowels were selected and further analyzed.

The first part of this thesis branches into two chapters where the topic is framed. In chapter one, the Portuguese European vowel system (PE) is described according to its standard variety, and in chapter two the relations between the back vowels and their orthographic representations are presented.

In the second part of the thesis, devoted to the empirical study, the relevant spelling mistakes are analyzed, according to the typology that was established to classify them, and also taking into consideration the learning level of PL2 of the informants.

To conclude, the causes that led the learners to produce such mistakes are discussed, in the attempt of contributing to the development of remedial strategies for learners of PL2.

# Índice

|   |     |
|---|-----|
| Agradecimentos  | i   |
| Resumo  | ii  |
| Abstract  | iii |
| Índice de figuras tabelas e gráficos                                | v   |
| Introdução  | 1   |
| Parte I: Breve enquadramento do tema                                | 2   |
| 1. Estrutura do sistema fonológico vocálico do português            | 3   |
| 2. Processos fonológicos: do vocalismo tónico<br>ao vocalismo átono | 6   |
| 3. Representação ortográfica das vogais posteriores<br>do português | 13  |
| Parte II: Trabalho empírico   | 18  |
| 1. Hipóteses  | 19  |
| 2. Metodologia de investigação                                      | 20  |
| 2.1 Caracterização dos informantes ( <i>subcorpus</i> )             | 21  |
| 2.2 Natureza dos dados  | 21  |
| 2.3 Procedimentos   | 23  |
| 3. Resultados e discussão   | 27  |
| 4. Conclusão  | 51  |
| Bibliografia  | 53  |
| Anexo 1 (em CD)   |     |
| Anexo 2 (em CD)   |     |

# Índice de figuras, tabelas e gráficos

## Figuras

|  |    |
|--|----|
| Figura 1: Processo do vocalismo átono do PE          | 8  |
| Figura 2.1: Regra das Vogais átonas [+rec]           | 9  |
| Figura 2.2: Regra das Vogais átonas [-rec]           | 9  |
| Figura 3: Regra de supressão de [i] em posição átona | 11 |

## Tabelas

|   |    |
|---|----|
| Tabela 1: Matriz binária das unidades vocálicas do PE   | 6  |
| Tabela 2: Relação fone-representação gráfica  | 15 |
| Tabela 3: Relação representação gráfica-fone  | 16 |
| Tabela 4: Categorias dos erros segmentais de acordo com a posição acentual  | 25 |
| Tabela 5: Categorias dos erros acentuais  | 26 |
| Tabela 6: Dez erros ortográficos mais frequentes, envolvendo a representação de vogais posteriores, por ordem decrescente, nos textos do " <i>Corpus PEAPL2</i> " | 33 |

## Gráficos

|  |    |
|--|----|
| Gráfico 1: Percentagem de palavras com erros ortográficos no universo de palavras do <i>corpus PEAPL2</i>                            | 28 |
| Gráfico 2: Percentagem de palavras com erros envolvendo vogal posterior relativamente ao universo de palavras com erros ortográficos | 29 |

|  |    |
|--|----|
| Gráfico 3: Percentagem de erros de representação segmental por posição acentual  | 31 |
| Gráfico 3.1: Percentagem de erros segmentais em posição tónica de acordo com a categoria   | 34 |
| Gráfico 4: Percentagem de erros por categoria  | 36 |
| Gráfico 4.1: Percentagem de erros de categoria 1 (o > u) por posição acentual  | 38 |
| Gráfico 4.2: Percentagem de erros de categoria 2 (u > o) por posição acentual  | 39 |
| Gráfico 4.3: Percentagem de erros de categoria 3 (o > ou) por posição acentual   | 39 |
| Gráfico 4.4: Percentagem de erros de categoria 4 (u > ou) por posição acentual   | 40 |
| Gráfico 4.5: Percentagem de erros de categoria 5 (ou > o) por posição acentual   | 40 |
| Gráfico 4.6: Percentagem de erros de categoria 6 (ou > u) por posição acentual   | 41 |
| Gráfico 4.7: Percentagem de erros de categoria 7 ([aw]/[ew])   | 41 |
| Gráfico 4.8: Percentagem de erros de categoria 8 (acentuação)  | 42 |
| Gráfico 5: Percentagem de palavras com erros ortográficos no universo de palavras obtidas por nível de aprendizagem dos informantes                                | 45 |
| Gráfico 5.1: Percentagem de erros ortográficos envolvendo vogal posterior no universo de palavras obtidas por nível de aprendizagem dos informantes                | 46 |
| Gráfico 5.2: Percentagem de erros ortográficos envolvendo vogal posterior no universo de palavras com erros ortográficos por nível de aprendizagem dos informantes | 46 |



# Introdução

A presente dissertação visa analisar a forma como as vogais da série posterior do Português Europeu são registadas nas produções escritas de aprendentes que têm o Português como língua não materna (PLNM)<sup>1</sup>. As peculiaridades do sistema vocálico do Português e, particularmente, das vogais posteriores, e a forma como os aprendentes representam estas últimas vogais foi, então, a motivação para o desenvolvimento deste trabalho.

Este estudo encontra-se dividido em duas partes. A primeira parte diz respeito ao sistema fonológico do português, e aí são descritas as características do seu sistema vocálico, apresentados os principais processos fonológicos, nomeadamente o processo do vocalismo átono (processo com implicações na aquisição do português e na aprendizagem da sua ortografia) e analisadas as correspondências que, no domínio das vogais posteriores, se estabelece entre o plano fonológico e o plano gráfico.

A segunda parte é dedicada ao tratamento dos dados. Tendo por base um *corpus* de produções escritas de aprendentes do português enquanto língua não materna, analisou-se esse *corpus* e, a partir daí, procedeu-se a uma selecção dos textos relevantes para o presente estudo. Deste modo, foi possível identificar inúmeras ocorrências de erros envolvendo vogais posteriores, categorizá-los de acordo com a natureza do desvio gráfico e a posição acentual em que ocorre para, assim, ser possível analisá-los e tirar conclusões sobre a forma como a série posterior das vogais é representada na produção escrita dos aprendentes estrangeiros.

---

<sup>1</sup> No âmbito deste trabalho, Língua Não Materna (LNM) é qualquer língua aprendida posteriormente à LM. O termo L2 pode também ser tomado nesta mesma acepção, pelo que se utilizarão LNM e L2 como equivalentes neste trabalho. Sobre o esclarecimento destes conceitos e de outros afins ver Isabel Leiria (2004).

---

# Parte I

---

Breve  
enquadramento  
do tema

---

## 1. Estrutura do sistema fonológico vocálico do português

O sistema fonológico do Português Europeu (PE) é composto por um conjunto de dezanove unidades consonânticas: (/p b t d k g m n ɲ f v s z ʃ ʒ l ʎ r R/) e por um inventário vocálico constituído por sete unidades /i e ε a ɔ o u/ (Mateus, 2006a: 1001, Câmara Jr., 1996: 41), sendo que quer consoantes, quer vogais apresentam variantes contextuais.

Embora uma parte das línguas do mundo recorra somente a cinco unidades vocálicas (/i, e, a, o, u/) (Maddieson, 2009: 126), que estão sempre presentes mesmo quando as línguas possuem um inventário vocálico maior<sup>2</sup>, o sistema de vogais do português é mais complexo, sendo, então, e como já se afirmou, constituído por /i/, /e/, /ε/, /a/, /ɔ/, /o/, /u/. O estatuto fonológico destes segmentos pode ser confirmado através de pares mínimos, isto é, de um conjunto de palavras que, variando apenas num segmento fónico, vêem o seu significado alterado como em *c[ɔ]me/c[u]me*, *p[e]lo/p[u]lo*, *s[a]lto/s[o]lto*, *p[ε]so/p[ε]so/p[i]so*, *m[ɔ]ro/m[u]ro/m[i]ro*. Estes conjuntos de palavras ilustram a pertinência fonológica das diferenças de timbre das vogais em questão e provam a existência, em português, das oposições fonológicas entre /ɔ/-/u/, /e/-/u/, /a/-/o/, /e/-/ε/-/i/, /ɔ/-/u/-/i/. Algumas destas oposições são muito rentáveis (casos de /ε/-/ɔ/ e /e/-/a/), enquanto outras já o serão menos (Mateus, 2006a: 991, 992), como é o caso das oposições /ε/-/e/, /ɔ/-/o/; evidencia-o a variação entre vogais baixas e médias (*dez[o]ito/dez[ɔ]ito*), de acordo com os hábitos regionais ou pessoais, mas também os poucos casos existentes de pares mínimos em que se verificam os contrastes em causa, *b[ɔ]la/b[o]la*, *s[ε]de/s[e]de*. Não invalidam esta informação os casos como *j[ɔ]go/j[o]go*

---

<sup>2</sup> Os sistemas fonológicos variam de língua para língua. Assim, há sistemas vocálicos com maior número de unidades, como o francês, e sistemas com menor número de unidades, como é o caso do espanhol e de algumas línguas caucásicas (Mateus, 2001: 2).

ou s[ε]lo/s[e]lo, uma vez que, aí, a oposição entre as vogais tónicas resulta da regra de harmonização vocálica que afecta as formas verbais. (Mateus, 2006b: 11 e 12).

A estrutura interna dos segmentos fonológicos vocálicos define-se pela presença de traços distintivos, ou seja, por um conjunto específico de propriedades identificadoras que constituem os segmentos fonológicos e que os opõem entre si. Uma teoria dos traços distintivos, que dá sequência a propostas interpretativas anteriores, foi desenvolvida, em 1968, por Chomsky e Halle, na obra *The Sound Pattern of English*, dando lugar a um sistema de classificação dos fonemas baseada em oposições binárias. Neste sistema, as unidades fónicas são identificadas através da presença (+) ou ausência (-) de determinada característica em relação a uma posição neutra<sup>3</sup>.

No português, segundo Maria Helena Mateus (2005: 187, 188), todas as unidades vocálicas são [-cons], ou seja, caracterizam-se por uma passagem livre do ar no tracto vocal, sem qualquer oclusão na cavidade bucal, dando-se uma vibração espontânea das cordas vocais, razão pela qual são também [+ soan] e, logo assim, necessariamente, [+ voz]<sup>4</sup>. São, para além disso, unidades [+ sil], na medida em que sozinhas, preenchem o

---

<sup>3</sup> Todos os traços apresentados nesta proposta de classificação se definem tomando como referência uma posição considerada "neutra". Chomsky e Halle (1968: 300) definem essa posição nos seguintes termos: " during quiet breathing (...) the velum is lowered, thereby allowing air to pass through the nose; in the neutral position, on the other hand, the velum is raised, and the air flow through the nose is shut off. The body of the tongue, which in quiet breathing lies in a relaxed state on the floor of the mouth, is raised in the neutral position to about the level that it occupies in the articulation of the English vowel [e] in the word bed. (...) the air pressure in the lungs just prior to speaking must be higher than the atmospheric pressure. During quiet breathing, the vocal cords must be widely spread apart since practically no sound is emitted. On the other hand, there is good reason to believe that prior to speaking the subject normally narrows his glottis and positions his vocal cords so that in the neutral position they will vibrate spontaneously in response to the normal, unimpeded air flow."

<sup>4</sup> O traço **soante** caracteriza-se por uma passagem do ar relativamente livre, originando uma vibração espontânea das cordas vocais. O traço **vozeado** indica a presença ou ausência da vibração das cordas vocais mas esta vibração pode ser espontânea ([+ soan]) ou voluntária ([- soan]). (Mateus, 2006a: 998)

núcleo silábico. Para além destes traços, que definem as vogais enquanto classe natural<sup>5</sup>, há que considerar, para a sua classificação, traços relacionados com a altura e localização da língua e com a existência ou não de ressonância labial.

Assim, enquanto na classificação tradicional as vogais se organizam em classes de acordo com a zona de articulação, isto é, em palatais, centrais e velares<sup>6</sup>, no sistema binário, e atendendo aos mesmos pormenores articulatorios, as vogais palatais /i, e, ε/ classificam-se como [- rec] e a central /a/ e as vogais velares /ɔ, o, u/, como [+ rec].

Quanto à participação do ressoador labial, só as vogais /ɔ, o, u/ são positivamente marcadas, logo [+ arr]. Não existem no PE padrão vogais centrais ou anteriores produzidas com ressonância labial e, logo assim, todas as centrais ou anteriores são [- arr].

Além da posição mais ou menos recuada do dorso da língua na cavidade bucal e da posição dos lábios, as vogais distinguem-se ainda pelo grau de elevação assumido pela língua aquando da sua articulação: podem ser baixas, médias ou altas. No que concerne então ao contraste de altura das vogais, patente em todas as línguas, /i, u/ são vogais altas (na sua articulação, o dorso da língua eleva-se na cavidade oral, porém sem criar uma obstrução que dê origem a uma consoante), enquanto /a, ε, ɔ/ são vogais baixas (implicam o abaixamento do dorso da língua em relação à sua posição neutra). Já as vogais /e, o/ são, por seu lado, [- alt] e [- bɿ].

Importa, a este propósito, salientar que o traço de altura, sendo um dos três parâmetros básicos para distinguir as vogais entre si, é válido em todas as línguas conhecidas. Assim, o parâmetro de altura das vogais parece ser a sua característica principal, pois

---

<sup>5</sup> Para que um conjunto de segmentos fonológicos constitua uma classe natural, é necessário que sejam caracterizados por um conjunto de traços distintivos comuns menor do que o necessário para especificar cada segmento. (Mateus, 1996:185, 186).

<sup>6</sup> Por seu turno, as consoantes organizam-se em classes de acordo com o ponto de articulação (no caso do português, labiais, dentais e alveolares, palatais, velares e uvulares), e o modo de articulação (oclusivas, fricativas, laterais, vibrantes e nasais).

todas as línguas o usam para estabelecer contrastes entre as vogais e nenhum outro parâmetro, nem o de localização, nem o de arredondamento, é usado universalmente. (Maddieson, 2009: 123).

A seguinte tabela sintetiza o que foi dito até agora quanto à caracterização dos segmentos vocálicos do Português Europeu (PE).

| Segmentos \ Traços | /i/ | /e/ | /ɛ/ | /a/ | /ɔ/ | /o/ | /u/ |
|--------------------|-----|-----|-----|-----|-----|-----|-----|
| [alto]             | +   | -   | -   | -   | -   | -   | +   |
| [baixo]            | -   | -   | +   | +   | +   | -   | -   |
| [recuado]          | -   | -   | -   | +   | +   | +   | +   |
| [arredondado]      | -   | -   | -   | -   | +   | +   | +   |

Tabela 1: Matriz binária das unidades vocálicas do PE, adaptado de Mateus e Andrade, 2000:30

Nesta tabela é possível identificar classes naturais de segmentos fonológicos, que possuindo traços comuns, estão sujeitos aos mesmos processos fonológicos (Mateus, 2006b:12, 13; Mateus, 2002: 7-9).

## 2. Processos fonológicos: do vocalismo tónico ao vocalismo átono

O Português é uma língua na qual actuam vários processos fonológicos, nomeadamente o da atribuição do acento principal de palavra, a assimilação e, um dos mais característicos na sua variedade europeia, o processo de elevação e recuo do vocalismo

átono, constituído por regras gerais mas também por restrições à sua ocorrência (como seguidamente será exposto).

Como se referiu na secção anterior, os processos fonológicos afectam classes naturais. O processo de redução do vocalismo átono, uma das peculiaridades do Português Europeu, consiste na alteração das vogais [- alt]<sup>7</sup>, quando em posição não acentuada.

Daí resulta que as sete vogais do PE (/i, e, ε, a, ɔ, o, u/), apresentando as realizações fonéticas [i, e, ε, a, ɐ<sup>8</sup>, ɔ, o, u] em posição tónica, registem apenas, quando átonas, os seguintes fones: [i, ɪ, ɐ, u]. Esta situação resulta, fundamentalmente, como veremos, da neutralização das diferenças entre vogais médias e baixas da mesma série de localização. Assim, o processo de redução do vocalismo átono diz respeito a uma alteração regular nas vogais em sílaba não acentuada<sup>9</sup> e traduz-se, em todos os casos, por alterações na altura das vogais afectadas: dá-se uma elevação das mesmas a qual atinge unidades médias e unidades baixas, alterando os seus traços de altura. Deste modo, devido à redução das vogais, [i] e [u] [+ alt] são o resultado das neutralizações das oposições entre /ε/ e /e/, [-alt] e [-rec] e entre /ɔ/, /o/, /u/, [+arr], respectivamente. Quanto a /a/, [+ bx], esta vogal sofre igualmente um fenómeno de elevação, materializando-se como [ɐ] [- bx] no mesmo contexto prosódico. Assim sendo, e dada a natureza essencialmente fonológica da ortografia portuguesa (Mateus, 2006b), assunto a que regressaremos adiante, nos casos em que este processo fonológico actua, um <o>

---

<sup>7</sup> É ainda relevante o facto de se verificar também um recuo de /i/, como se pode comprovar nas formas verbais em que esta vogal se situa em final átono de palavra como em *foge* (versus *fugir*), *sente* (versus *sentir*).

<sup>8</sup> Esta unidade é, também em posição tónica, o resultado de um processo fonológico de elevação.

<sup>9</sup> Este é um processo de grande relevância a nível fonológico e acessível a qualquer falante da língua, isto é, qualquer falante de português aplica a regra de redução das vogais, sem que para isso faça qualquer esforço consciente.

gráfico em posição átona, em vez de se pronunciar [ɔ] ou [o], pronuncia-se habitualmente [u] e um <e> gráfico, igualmente átono, não se lê [e] nem [ɛ], mas sim, previsivelmente, [i].

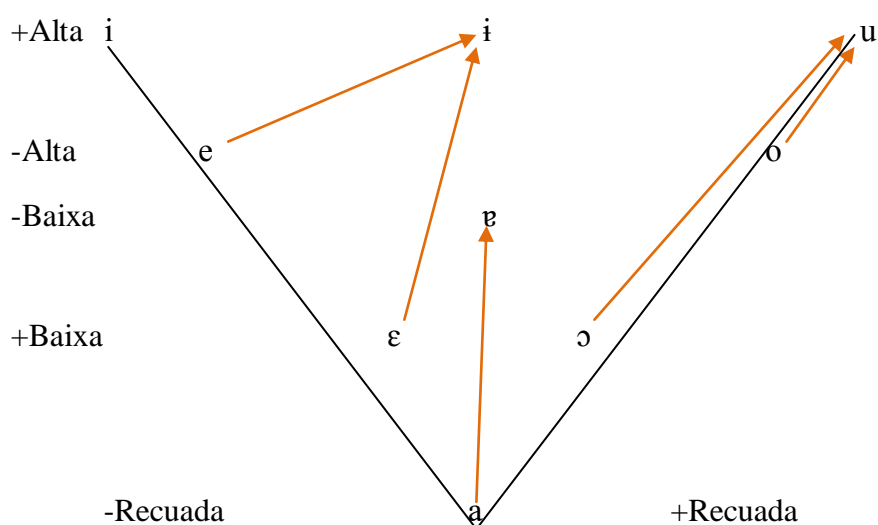


Figura 1: Processo do vocalismo átono do PE  
(adaptado de Mateus, 2006b:12)

Através da figura apresentada, verifica-se que o vocalismo átono do português está sujeito a um processo, não só de elevação, mas igualmente de recuo, que afecta as vogais /e/ e /ɛ/. Estas, quando em posição átona, convergem em [i]: *m[e]sa/ m[i]sinha*, *b[ɛ]la/ b[i]leza*. Também as vogais [+rec] convergem em [u]: *c[o]rro/ c[u]rrida*, *p[ɔ]rta/ p[u]rteiro*, *soc[o]rro/ soc[u]rrer*, *aut[ɔ]grafo/ aut[u]grafado*.

As transformações descritas (de /e/ e /ɛ/ [-alt, -rec], e /o/ e /ɔ/ [-alt, +rec, +arr]) podem ser representadas sob a forma das seguintes regras fonológicas.



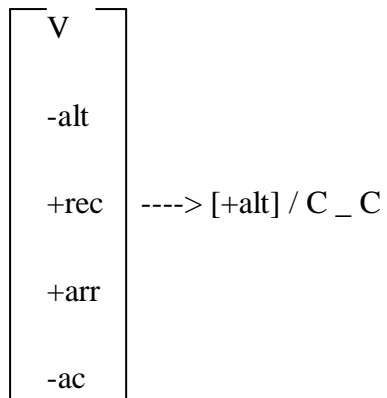


Figura 2.1: Regra das Vogais átonas [+rec]  
(Mateus, 1996: 191)

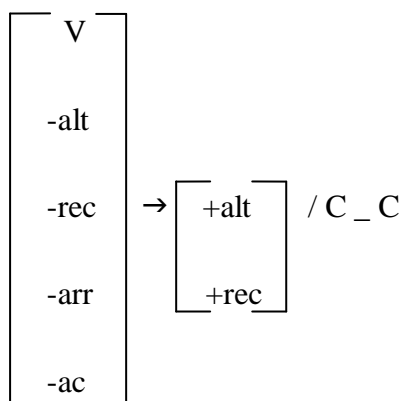


Figura 2.2: Regra das Vogais átonas [-rec]  
(Mateus, 1996: 191)

Assim, ao contrário do que sucede com as vogais arredondadas, que sofrem um processo de elevação convergindo em [u] em sílaba átona, as vogais [-recuadas] não convergem com a vogal [+alt] da mesma série, isto é, com [i], exceptuando-se o que acontece em posição átona final nas formas verbais.

Da actuação dos processos assim formalizados resulta o desaparecimento das vogais baixas, isto é, das vogais que, em posição tónica, solicitam um abaixamento da língua, /ɛ/, /ɔ/<sup>10</sup>, surgindo, por outro lado, um timbre que não tem ocorrência em posição tónica, [i]. Esta vogal, que só ocorre em sílabas átonas, encontra-se, portanto, em distribuição complementar com [e] e [ɛ] tónicos, sendo um alofone (isto é, uma das diferentes realizações fonéticas possíveis) dos fonemas ou segmentos /e/ e /ɛ/.

É também relevante referir que, em posição átona final, apenas existem três unidades [i, ɐ, u] e que, relativamente a [i], é recorrente o seu desaparecimento na fala coloquial (essa unidade só ocorre em registo pausado no PE).

Apesar da regularidade das alterações provocadas pelo processo de elevação pelo qual as vogais átonas passam, há determinados contextos em que ele não actua. Assim sendo, verificam-se dois tipos de excepções à actuação do processo: (i) as excepções regulares, determinadas por restrições previsíveis, e (ii) as excepções marcadas no léxico. As primeiras caracterizam-se pelo facto de ser possível, através do contexto, identificar o que impede a aplicação das regras do vocalismo átono; as segundas, pelo contrário, distinguem-se por existir um impedimento à aplicação das regras que não é identificável pelo contexto, ficando tais casos marcados a nível lexical (Mateus, 2006a: 1013-1016).

No que concerne às excepções regulares, os contextos que impossibilitam a aplicação das regras do vocalismo átono são os seguintes:

(i) o facto de as vogais átonas pertencerem a núcleos de sílabas com /l/ a preencher a coda, como é o caso em *g[ɔ]lpe-g[ɔ]lpear*, *r[ɛ]lva-r[ɛ]lvad*, *m[a]l-m[a]lvado*;

(ii) as vogais fazerem parte de ditongos decrescentes: *d[e]us-end[e]usar*, *pn[e]u-pn[e]umático*, *b[a]irro-b[a]irrista*;

(iii) as vogais encontrarem-se em início absoluto de palavra (*h[ɔ]spital*, *[o]lhar*).

---

<sup>10</sup> O mesmo acontece a /a/, sujeito a um processo de elevação que se pode representar de forma análoga à que se encontra nas Figuras 2.1. e 2.2.

As palavras derivadas por sufixos -z avaliativos ou sufixo -mente, como *s[ɔ]zinha*, *s[ɔ]mente*, são também excepções às regras gerais que afectam as vogais átonas (Mateus, 2006a: 1013-1015).

No caso das excepções marcadas no léxico (como *abs[ɔ]rver*, *ad[ɔ]ptar*, *direct[o]r*, *red[a]cção*) actuaram processos históricos cujas regras o falante não consegue inferir. Por este motivo, os timbres das vogais em questão têm de ser aprendidos e memorizados pelos falantes (Mateus, 2006a: 1015, 1016).

A alteração resultante no timbre das vogais é um dos aspectos essenciais associados ao processo de acentuação do português e o processo de redução vem nalguns casos acompanhada de supressão das vogais a nível fonético que, por outro lado, apresenta diferenças em função da posição átona em causa. Assim, as vogais átonas das sílabas finais estão mais sujeitas à supressão que as pré-tónicas; a mais vulnerável a este processo é a vogal átona [i], que apenas ocorre em registo pausado sendo, não raras vezes, suprimida no registo coloquial: *menina* [m'ninɐ], *teste de* ['tɛst d].

A supressão da vogal átona [i] pode ser representada formalmente através de uma regra:

$$\left[ \begin{array}{c} \text{V} \\ +\text{alt} \\ \\ +\text{rec} \\ \\ -\text{arr} \end{array} \right] \rightarrow \emptyset$$

Figura 3: Regra de supressão de [i] em posição átona (Mateus., 1996: 192)

Relativamente à vogal central /a/, que, em posição átona, também se eleva, realizando-se como média [ɐ], o facto de não se tornar [+ alt], faz com que essa unidade resista melhor à supressão.

A redução e frequente apagamento das vogais átonas do PE em posição pós-tónica e em posição átona final (embora estes fenómenos possam também ocorrer em sílaba pré-tónica) afectam não só [i], mas também [u] (em posição final de palavra: *estudo*-[ʃtúd] ou precedido de uma labial: *fonologia*-[fnulu'ziɐ]), apesar de esta supressão ser mais frequente no caso da primeira vogal. As mesmas vogais, [i] e [u], quando, em posição átona, seguidas de outra vogal (como em *cear* e *soar*), realizam-se como semivogais e, neste caso, nunca podem ser suprimidas. Por outro lado, quando há encontro das vogais átonas [ɐ], [i] e [u] em final de palavra com as vogais iniciais da palavra seguinte, dá-se a supressão da átona final com uma consequente simplificação das duas sílabas em contacto numa única, como é ilustrado pelos exemplos *cama horrível* - ['kɐmɔ'rivɛɫ], *trouxeste água* - [tro'sɛʃt'agwɐ] (Mateus, informação consultada em: [http://cvc.instituto-camoes.pt/cpp2/acessibilidade/capitulo9\\_1.html](http://cvc.instituto-camoes.pt/cpp2/acessibilidade/capitulo9_1.html)).

No caso da vogal átona [i], outro contexto de supressão é o início de palavra quando seguida de [ʃ] ou [ʒ] (como em *estar* ['ʃtar], *esganar* [ʒgɐ'nar]).

Já a vogal final [ɐ] sofre crase quando a palavra seguinte começa por [ɐ] ou [a], daí resultando [a] (*casa alegre* - ['kaz'alɛgr], *da altura* - [d'aɫturɐ]).

Finalmente, no caso de vogal final átona [u] seguida por palavra iniciada por uma vogal baixa [a, ɔ, ɛ], por vogal média [e], e por vogal alta [i], dá-se uma semivocalização ([wa], [wɔ], [wɛ], [we], [wi], *carro alto* - ['kar'waɫtu], *tanto ódio* - ['tɛ̃t'wɔdiu], *como era* - ['kom'wɛrɐ], *como este* - ['kom'wɛʃti], *tudo isto* - ['tud'wiʃtu], respectivamente). Tal não ocorre no caso de palavra iniciada por [o]; desse encontro resulta [o] (*muito orgulho* - ['mujtɔr'guɫu]).

Importa ainda referir que, nas palavras terminadas em [i] átono e seguidas por outra palavra iniciada por vogal, também o [i] semivocaliza (*júri atento* - ['ʒurjɐ'tɛ̃tu]). (Mateus, informação consultada em: [http://cvc.instituto-camoes.pt/cpp2/acessibilidade/capitulo3\\_2.html](http://cvc.instituto-camoes.pt/cpp2/acessibilidade/capitulo3_2.html)).

A redução vocálica não se manifesta da mesma forma em todas as vogais, isto é, embora todas elas sofram um processo de elevação subindo um ou dois graus em relação à sua configuração em posição tónica, nem todas sofrem um processo de recuo como o verificado na série anterior das vogais do PE. Em síntese, no processo do vocalismo átono, a redução manifesta-se, então, no facto de as vogais se elevarem quando não acentuadas. Por sua vez, o recuo diz respeito às vogais fonológicas [-rec], pois /e/ e /ɛ/ têm como realização [i], [+rec].

De acordo com Maria Helena Mateus (2006b: 11), as vogais fonológicas, com várias correspondências fonéticas, estabelecem uma complexa relação com as suas representações no plano gráfico, já que os grafemas <a>, <e>, <i>, <o> e <u> abrangem mais do que uma realização. Assim sendo, <a> representa os fones [a] e [ɐ], <e>, a vogal com mais correspondências, regista [ɛ, e, ɐ, i, i, j], <i> reproduz [i] e [j]<sup>11</sup>, <o> representa [ɔ, o, u, w] e, por fim, <u> representa [u] e [w].

### **3. Representação ortográfica das vogais posteriores do português**

Enquanto a variação geográfica social e situacionalmente motivada manifestada na oralidade é uma propriedade inerente a qualquer língua, não dando lugar, numa perspectiva descritiva, a valorações, na ortografia é possível aplicar a oposição correcto/incorrecto na análise de formas alternativas. O erro ortográfico não é nem social, nem academicamente tolerado, mas a aprendizagem da ortografia enfrenta alguns desafios relevantes.

Porventura o mais evidente destes desafios decorre do facto de, e como referido na secção anterior, não existirem relações biunívocas entre as realizações fónicas dos segmentos fonológicos e as respectivas representações ortográficas. Esta será a

---

<sup>11</sup> Nalguns casos, <i> pode igualmente representar [i], como em *m[i]nistro*.

problemática a desenvolver nesta secção, onde nos centraremos nas representações gráficas das unidades da série posterior.

Da descrição já empreendida, pode concluir-se que este desfasamento entre unidades dos planos fónico e gráfico atinge particular expressividade no caso das unidades desta série. Como se viu, da actuação do processo de elevação em sílaba átona, geral no PE, das unidades média e baixa resulta o timbre [u], sendo aquelas, no entanto, graficamente representadas como <o><sup>12</sup>. Contudo, para [u], o sistema ortográfico do português prevê a representação <u>.

Poder-se-á admitir que estes serão factores que dificultam a aprendizagem da ortografia e induzem ao erro ortográfico, quer em aprendentes que têm o português como língua materna, quer junto de aprendentes de português como língua não materna (PLNM).

Mas outros processos, que podemos descrever como representativos de variação geográfica e diacrónica, podem ser igualmente fonte de dificuldades na aprendizagem da ortografia. Referimo-nos, concretamente, à crescente tendência para se pronunciar de forma monotongada o ditongo grafado <ou>. Na sua forma plena, esse ditongo preserva-se apenas no norte de Portugal e em grande parte dos dialectos brasileiros (Mateus, 2005: 165, 166).

Admitindo que em determinadas fases de aprendizagem da ortografia, a oralidade exerce uma influência determinante sobre a escrita, será importante ter em consideração o pormenor das relações fone-grafema e grafema-fone no PE.

Nas tabelas seguintes estão esquematizadas as relações **fone-representação gráfica** e **representação gráfica-fone** no PE, no que diz respeito a unidades da série posterior. Na primeira tabela, encontram-se as representações gráficas das vogais (orais e nasais) da série posterior do português, e, na segunda, as representações fonéticas que cada uma das representações gráficas concretiza.

---

<sup>12</sup> O facto de a demarcação entre as vogais baixa e média, [ɔ] / [o], não estar representada por símbolos gráficos diferentes reflecte, de alguma forma, a pouca produtividade fonológica desta diferença de timbres, já assinalada na secção anterior.

| Fone | Contexto                           | Exemplo                            | Representação gráfica |
|------|------------------------------------|------------------------------------|-----------------------|
| [u]  | Posição átona pré-tónica           | Monu <b>mento</b>                  | <u>                   |
|      |                                    | Abor <b>recido</b>                 | <o>                   |
|      | Posição tónica                     | Descu <b>bro</b>                   | <u>                   |
|      |                                    | Ú <b>ltimo</b>                     | <ú>                   |
|      | Posição átona final                | Cabe <b>lo</b>                     | <o>                   |
|      | Posição átona (clíticos)           | No                                 | <o>                   |
| [ũ]  | Posição tónica ou átona + clíticos | Mun <b>do</b> /mundial             | <un>                  |
|      |                                    | Um / at <b>um</b>                  | <um>                  |
| [õ]  | Posição tónica ou átona            | Bom/Compu <b>ta</b> dor            | <om>                  |
|      |                                    | Pon <b>te</b> /Con <b>stru</b> ção | <on>                  |
| [o]  | Posição tónica                     | Po <b>vo</b>                       | <o>                   |
|      |                                    | Av <b>ô</b>                        | <ô>                   |
|      |                                    | Po <b>ucos</b>                     | <ou>                  |
|      | Posição átona + clíticos           | <b>Ou</b> /dourar                  | <ou>                  |
|      | Posição pré-tónica                 | Sold <b>ado</b>                    | <o>                   |
| [ɔ]  | Posição átona pré-tónica           | So <b>zinha</b>                    | <o>                   |
|      | Posição tónica                     | Sort <b>e</b>                      | <o>                   |
|      |                                    | Domin <b>ó</b>                     | <ó>                   |

|     |                               |          |     |
|-----|-------------------------------|----------|-----|
| [w] | Glide em ditongo (decrecente) | Bacalhau | <u> |
|     | Glide em ditongo (crescente)  | Soar     | <o> |
|     |                               | Ruela    | <u> |

Tabela 2: Relação fone-representação gráfica

| Representação gráfica | Contexto                 | Exemplo  | Fone            |
|-----------------------|--------------------------|----------|-----------------|
| <u>                   | Posição átona pré-tónica | Pular    | [u]             |
|                       | Posição tónica           | Lula     | [u]             |
|                       | Ditongo (decrecente)     | Bacalhau | [w]             |
|                       | Ditongo (crescente)      | Luar     | [w]             |
| <ú>                   | Posição tónica           | Úvula    | [u]             |
| <o>                   | Posição átona pré-tónica | Operário | [ɔ] / [o] / [u] |
|                       |                          | Moer     | [w]             |
|                       |                          | Portão   | [u]             |
|                       | Posição tónica           | Porta    | [ɔ]             |
|                       |                          | Novo     | [o]             |
|                       | Posição átona final      | Pêlo     | [u]             |



|      |                         |                 |     |
|------|-------------------------|-----------------|-----|
| <ó>  | Posição tónica          | Avó             | [o] |
| <ô>  | Posição tónica          | Avô             | [o] |
| <ou> | Posição tónica ou átona | <b>Ou</b> rives | [o] |
|      |                         | <b>Ou</b> ro    |     |
|      |                         | Cant <u>ou</u>  |     |

Tabela 3: Relação representação gráfica-fone

Todos os aspectos enunciados e, particularmente, o facto de, nas vogais da série posterior, a elevação fazer coincidir um fone com outro para o qual estaria previsto outro grafema (em posição átona <o> actualiza-se como [u] e <u> concretiza-se igualmente como [u]) tornam esta série particularmente interessante e constituem a grande motivação deste trabalho. Tendo em conta estas especificidades foram analisados os textos pertencentes ao *Corpus* de Produções PEAPL2 e estudou-se de que forma se reflectiam, na aprendizagem da escrita, essas características da série posterior das vogais. O resultado dessa análise será então apresentado seguidamente neste trabalho.

---

# Parte II

---

Trabalho empírico

---

## 1. Hipóteses

Após a análise de todos os textos pertencentes ao *Corpus* de Produções Escritas de Aprendentes de PL2 (PEAPL2), foi possível colocar algumas hipóteses tendo em conta:

- i) as características do sistema fonológico vocálico do português, concretamente das vogais da série posterior e dos processos fonológicos que sobre elas actuam;
- ii) a natureza das relações entre o sistema fonológico e a ortografia do português;
- iii) o tipo de aprendente que forneceu os dados a analisar neste trabalho.

Assim, as hipóteses que vamos equacionar serão verificadas através da análise dos dados empíricos a apresentar nesta parte do trabalho.

Considerando a existência de erros ortográficos recorrentes na representação de unidades vocálicas da série posterior junto de aprendentes de PLNM e dada a natureza fonológica da ortografia do português poderá prever-se que uma parte relevante desses erros ortográficos se ficará a dever:

- i) ao processo fonológico de redução vocálica, que afecta as vogais em posição átona e perturba a relação das unidades dos dois planos, gráfico e fónico;
- ii) a um eventual efeito regulador da oralidade sobre a escrita levando os aprendentes, e em especial os menos proficientes, a basear as suas opções gráficas na percepção dos modelos fónicos a que estão expostos.

Com a crescente exposição ao *input* gráfico proporcionado pela experiência de ensino em PL2, a hipersensibilidade aos padrões fónicos deverá, pois, diminuir progressivamente dando lugar à menor incidência de erros ortográficos nos níveis de aprendizagem mais elevados.

## 2. Metodologia de investigação

Este estudo parte da análise do modo como aprendentes de Português Língua Não Materna (PLNM) registam, na escrita, as unidades vocálicas posteriores. Para tal, tivemos em consideração as ocorrências desses segmentos quer em sílaba tónica, quer nas diversas posições átonas (recorde-se, para justificar esta opção, a natureza dos processos fonológicos próprios do Português Europeu que descrevemos no capítulo 1 da Parte I deste trabalho).

O *corpus* de textos escritos utilizado nesta investigação foi extraído do *Corpus* de Produções Escritas de Aprendentes de Português L2 (PEAPL2) do Centro de Estudos de Linguística Geral e Aplicada (CELGA) (disponível em <http://www.uc.pt/fluc/rcpl2>), que foi criado com o intuito de servir diversos interesses de investigação sobre Português L2 (PL2). Este *corpus*, que enriquece o conjunto de dados empíricos disponível aos investigadores de PL2, dada a forma como foi construído, permite observar, entre outros aspectos, as configurações das interlínguas dos informantes ao longo do processo de aprendizagem (Martins, 2011). A partir da análise dos textos do *corpus* PEAPL2, constituiu-se, então, um *subcorpus* específico que serviu de base à investigação para esta dissertação, sendo que esse *subcorpus* diz respeito aos textos em que se manifestaram erros ortográficos envolvendo as vogais posteriores. Doravante, identificar-se-á o subconjunto de textos utilizados na presente investigação com recurso à expressão "*subcorpus*".

Os textos escritos que analisámos foram produzidos por alunos que frequentaram a Universidade de Coimbra no período compreendido entre Maio de 2009 e Maio de 2010, e que estavam inscritos nos Cursos de Português para Estrangeiros da Faculdade de Letras.

A recolha dos textos foi feita durante as aulas do *Curso Anual de Língua e Cultura Portuguesas para Estrangeiros*, do *Curso de Férias de Língua e Cultura Portuguesas para Estrangeiros* e do curso de *Língua Portuguesa Erasmus*. Os informantes distribuem-se por diversos níveis de aprendizagem, definidos de acordo com o proposto no Quadro Europeu Comum de Referência para as Línguas (QECL). Assim,

colaboraram na recolha de *corpus* aprendentes que integravam os níveis A1, A2, B1, B2 e C1<sup>13</sup>.

## 2.1 CARACTERIZAÇÃO DOS INFORMANTES (*SUBCORPUS*)

Os 391 aprendentes de PL2 que, através dos seus textos, tornaram possível este trabalho de investigação provêm de diferentes países e culturas, variando as suas idades entre os 16 e os 68 anos. Destes, são 264 os autores dos textos do *subcorpus* seleccionado a partir do *Corpus* PEAPL2, sendo de ambos os sexos, representando 41 nacionalidades diferentes e falando 31 línguas maternas (LM) distintas. Importa ainda referir o facto de todos os informantes saberem outra L2, além do português, na maior parte dos casos o inglês, e de nem todos falarem o português fora do contexto escolar.

## 2.2 NATUREZA DOS DADOS

Este grupo de 264 informantes, bastante heterogéneo, produziu, então, o *subcorpus* de 320 textos, através do qual se pode, entre outras potencialidades, analisar a produção escrita tendo em conta os erros ortográficos que envolvem as vogais posteriores.

---

<sup>13</sup> Note-se que a determinação do nível do aprendente de acordo com o proposto no QECRL não tem relação directa com o seu grau de escolaridade formal. Aos informantes foi atribuída uma classificação de nível de proficiência com base no na turma/nível/tipo de curso frequentado na FLUC. As limitações deste procedimento foram claramente identificadas em Martins (2011) onde se escreve o seguinte: “a forma como se distribuem os alunos pelos distintos **níveis/turmas** nem sempre reflecte de modo fiel os efectivos níveis de progressão linguística. Numerosos factores, alheios à organização dos cursos, contribuem para tal circunstância, que se faz sentir, mais particularmente, nas unidades curriculares de *Língua Portuguesa Erasmus*. Neste sentido, pode ser útil complementar as informações relativas ao nível/turma com os resultados da auto-avaliação”.

Os textos centram-se em três áreas temáticas: *o indivíduo, a sociedade e o meio ambiente*. Para o seu desenvolvimento, foram apresentados vários estímulos (três para cada um dos temas), adequados a cada nível de competência/capacidade, e que os aprendentes puderam seleccionar, na maioria dos casos, livremente. Deste modo, na temática *o indivíduo*, foram propostos os seguintes estímulos:

- i) "Fale daquilo que gosta de fazer nos tempos livres";
- ii) "Escreva uma carta a um amigo que não vê há muito tempo. Recorde momentos passados em conjunto e fale-lhe da sua vida pessoal e profissional actual";
- iii) "Escreva um texto em que se apresente, em que fale das suas características físicas, da sua vida familiar, da sua casa, dos seus gostos e dos seus desejos. Se não quiser falar de si, pode inventar".

Já do tema *a sociedade* fazem parte os estímulos:

- i) "Certamente já teve oportunidade de contactar com pessoas de cultura diferente da sua. Fale de um episódio que lhe recorde esse momento, das dificuldades sentidas, das diferenças e semelhanças encontradas entre as duas culturas e das experiências que partilharam";
- ii) "Há, certamente, comidas de que gosta muito e há outras que detesta. Fale disto e daquilo que pensam os seus familiares e amigos sobre o assunto";
- iii) "Todos os países são diferentes a nível cultural e geográfico. Descreva o seu país, observando as particularidades das suas regiões, os principais monumentos e saliente alguns dos hábitos mais frequentes da sua cultura".

Finalmente, quanto ao tema *o meio ambiente*, o aluno podia desenvolver textos a partir dos seguintes estímulos:

- i) "Fale do bairro onde mora. Diga se gosta dele e se acha que há coisas que podiam mudar para que fosse mais agradável lá viver";
- ii) "Fale de meios de transporte. Fale daqueles em que já viajou e daqueles em que gostaria de viajar. Se quiser, pode contar uma viagem que tenha feito";

iii) "Gosta de viver na cidade? Acha que, se pudesse, gostaria mais de viver no campo? Pense em vantagens e desvantagens de viver na cidade ou no campo. Escreva sobre isso".

### 2.3 PROCEDIMENTOS

Após a análise da totalidade dos textos do *corpus* PEAPL2 (isto é, de 546), produzidos pelos aprendentes de diversos níveis (ao nível A1 e A1+ correspondem 111 textos, ao nível A2 e A2+ correspondem 117, do nível B1 foram produzidos 251 textos, do B2 produziram-se 43 e, finalmente, do nível C1 produziram-se 24), apurou-se um total de 6098 erros ortográficos, dos quais 814 dizem respeito a erros que envolvem a representação de vogais posteriores. Estes 814 erros podem ainda ser divididos em erros segmentais (370 erros) e erros acentuais (444 erros) (como se verificará seguidamente nas tabelas 4 e 5).

Os erros ortográficos patentes nas produções dos aprendentes foram então destacados a negrito, de forma a mais facilmente serem identificados, tendo sido posteriormente transpostos para um folha Excel<sup>14</sup>, elaborada a partir da matriz disponibilizada na página de divulgação do *Corpus* PEAPL2 (<http://www.uc.pt/fluc/rcpl2/dados/>) e na qual se encontram organizados, por secções, os dados caracterizadores dos informantes e por eles fornecidos através do seu auto-testemunho.

Esclareça-se que foram ainda recolhidos dados pessoais e informações sobre a auto-avaliação de cada informante relativamente à proficiência nas LNM conhecidas, incluindo aí o português<sup>15</sup>.

---

<sup>14</sup> Os dados assim organizados serão fornecidos em anexo (sob a forma de CD).

<sup>15</sup> Cada secção representa uma coluna na folha Excel referida. Na primeira, *identificação da recolha*, encontram-se somente informações que possibilitam a identificação do informante através de um código que lhe é atribuído em função do curso, da turma/nível e número individual. Os dados pessoais encontram-se na segunda secção, *Dados pessoais do informante* e aí se tem acesso à data de nascimento, sexo, nacionalidade e país em que o aluno nasceu e viveu. Os dados acerca do contacto que o aluno já teve com línguas surge na terceira secção,

Após as secções com informação relativa à caracterização dos informantes, organizámos, então, em colunas, os erros ortográficos, quer segmentais, quer acentuais que envolvem as unidades posteriores, que haviam sido detectados nos textos produzidos por cada um. Assim, foi criada uma coluna que corresponde ao excerto onde se encontra o erro (assinalado a negrito), outra coluna com o registo do erro e da palavra na sua forma gráfica correcta; na coluna seguinte encontra-se a classificação dos erros de acordo com a tipologia criada para o efeito (descrita e apresentada no seguimento desta secção). Finalmente, na última coluna regista-se o número de vezes que o mesmo erro se repete no mesmo texto. Deste modo, cada erro dado por determinado informante no mesmo texto corresponde a uma linha (com o respectivo número de vezes que se repete nessa produção), perfazendo assim um total de 694 linhas com erros envolvendo unidades da série posterior, sendo que a cada linha só corresponde um erro (que pode ocorrer no mesmo texto uma ou mais vezes), somando assim o total de 814 erros envolvendo vogais posteriores.

Assim, e de acordo com o que foi descrito, após a análise dos 6098 erros ortográficos segmentais e acentuais consideraram-se somente os erros que envolvem vogais posteriores nas suas diversas realizações (quer como [ɔ], [o], [u] ou [w]) e procedeu-se à sua classificação e contagem em função da respectiva categoria: da categoria 1 fazem parte os erros em que o aprendente utiliza <u> em vez de <o>; da categoria 2, os erros em que usa <o> em vez de <u>; da categoria 3, os erros em que usa <ou> em vez de <o>; da categoria 4, os erros em que utiliza <ou> em vez de <u>; da categoria 5 os erros em que o aprendente utiliza <o> em vez de <ou>; da categoria 6, os erros em que usa <u> em vez de <ou> e, finalmente, da categoria 7, os erros que envolvem o registo de [w] nos ditongos [aw] e [ew].

---

*Línguas*, da qual fazem parte informações relativas à língua materna, à língua de escolarização e ainda ao conhecimento de outras línguas não maternas. Segue-se a secção *LNM com maior proficiência*, onde se encontra a auto-avaliação do informante relativamente à sua proficiência nas diferentes vertentes de uso linguístico, quer na modalidade escrita, quer na oral. Na secção *Português* encontram-se informações relativas ao tempo de aprendizagem formal de português, aos contextos em que esta língua é utilizada fora da situação de aula e a auto-avaliação de cada aluno relativamente à sua competência na escrita e na oralidade.



Ainda no que diz respeito aos erros segmentais, foram tidas em consideração todas as posições prosódicas e os que ocorrem em clíticos (tal como já foi anteriormente referido e como se pode verificar na tabela 4).

No que concerne aos erros acentuais, também estes se dividem em 5 tipos envolvendo a utilização do diacrítico<sup>16</sup>. Assim, consideram-se cinco situações distintas: i) a atribuição de acento gráfico ou til a uma sílaba átona, por deslocação do diacrítico em palavras que o contenham; ii) a atribuição de acento gráfico ou til a uma sílaba átona em palavras que não têm registo gráfico do acento (incluindo clíticos); iii) a atribuição de acento gráfico ou til a uma sílaba tónica em situações não previstas na grafia oficial; iv) a omissão de acento gráfico ou til; v) a selecção do tipo incorrecto de diacrítico para assinalar o acento prosódico (estes dados encontram-se apresentados na tabela 5).

| CATEGORIA <sup>17</sup> | POSIÇÃO ACENTUAL                 | EXEMPLO                 |
|-------------------------|----------------------------------|-------------------------|
| 1. <o> > <u>            | 1.1 Posição pré-tónica           | Aburrecido (aborrecido) |
|                         | 1.2 Posição átona final          | Metu (meto)             |
|                         | 1.3 Posição tónica               | Muro (moro)             |
|                         | 1.4 Posição átona (clíticos)     | Nu (no)                 |
| 2. <u> > <o>            | 2.1 Posição pré-tónica           | Monommento (monumento)  |
|                         | 2.2 Posição tónica <sup>18</sup> | Sobretudo (sobretudo)   |
|                         | 2.3 Posição átona (clítico)      | Nom (num)               |
| 3. <o> > <ou>           | 3.1 Posição átona pré-tónica     | Journal (jornal)        |
|                         | 3.2 Posição átona final          | Passou (passo)          |
|                         | 3.3 Posição tónica               | Pouvo (povo)            |
|                         | 3.4 Posição átona (clítico)      | Pour (por)              |

<sup>16</sup> Associamos, aqui, os casos em que o til é impropriamente utilizado, não pela sua função na marcação de traço prosódico, mas por se tratar, como os acentos grave, agudo e circunflexo, de um sinal auxiliar de escrita.

<sup>17</sup> O sinal > indica que a forma gráfica prevista (à esquerda) é substituída pelo grafema apresentado à direita.

<sup>18</sup> Foram incluídos nesta subcategoria os casos únicos de *entroido* (*Entrudo*) e *escoito* (*escuto*) registados na produção escrita de um único informante galego. Nestas circunstâncias, considerou-se apenas a alteração gráfica da vogal e não o processo de ditongação concomitante.

|   |                             |                        |
|---|-----------------------------|------------------------|
| 4. <u> > <ou>   | 4.1 Posição pré-tónica      | Tourísticos (turistas) |
|   | 4.2 Posição tónica          | Cursos (cursos)        |
| 5. <ou> > <o>   | 5.1 Posição tónica          | Pocos (poucos)         |
|   | 5.2 Posição pré-tónica      | Robado (roubado)       |
|   | 5.3 Posição átona (clítico) | O (ou)                 |
| 6. <ou> > <u> <sup>19</sup>                                 | 6.1 Posição pré-tónica      | Lusã (Lousã)           |
|   | 6.2 Posição tónica          | Muodu (mudou)          |
| 7. Representação da semivogal [w] (excepto no ditongo <ou>) | 7.1 [aw] _ <ao>             | Bakalhao (bacalhau)    |
|   | 7.2 [ew] _ <eo>             | Meos (meus)            |

Tabela 4: Categorias dos erros segmentais de acordo com a posição acentual

| ACENTUAÇÃO  |   | EXEMPLO  |
|---|---|--|
| 8.1 Acento gráfico ou til sobre vogal ou semivogal átona                      | 8.1.1 por deslocação do acento ou do til em palavras que os contenham | Misericórdia (misericórdia)<br>Estereótipos (estereótipos) |
|   | 8.1.2 em palavras sem acento gráfico ou til (incluindo clíticos)      | Eufória (euforia)<br>Oú (ou)                               |
| 8.2 Acento gráfico ou til sobre vogal tónica (não previsto na grafia oficial) |   | Escóla (escola)<br>Pessoa (pessoa)                         |
|   | 8.3 Falta de acento gráfico ou til                                    | Ludicas (lúdicas)<br>Milhoes (milhões)                     |
| 8.4 Uso do tipo incorrecto de acento  |   | Avôs (avós)<br>Sò (só)                                     |

Tabela 5: Categorias dos erros acentuais

<sup>19</sup> Os exemplos apresentados correspondem aos únicos registos obtidos para esta categoria. Ainda assim, e uma vez que se poderá tratar de uma manifestação de redução do ditongo <ou>, optou-se por contemplar esta categoria na tipologia.

Foi a partir dos dados assim organizados que se elaboraram os gráficos apresentados e discutidos na secção 3 do presente capítulo dos quais faremos (na secção seguinte) a análise de pormenor, de modo a assim verificarmos a validade das hipóteses de trabalho que colocámos.

### **3. Resultados e discussão**

De acordo com o que foi referido anteriormente, calculou-se a percentagem de palavras com erros ortográficos no universo de palavras do *Corpus* PEAPL2 (gráfico 1), bem como a quantidade de palavras com erros envolvendo vogal posterior relativamente ao total de palavras com erros (gráfico 2). Procedeu-se também à observação da distribuição de erros segmentais por posição acentual (gráfico 3). Foram ainda calculadas as percentagens relativas: às diferentes categorias de erros segmentais representada na tabela 4 em posição tónica (gráfico 3.1), aos erros globais por cada categoria da mesma tipologia (gráfico 4) e aos erros, por categoria e posição acentual (gráficos 4.1 a 4.8).

Elaborou-se, por fim, um conjunto de três gráficos nos quais os dados foram organizados tendo em conta o nível de aprendizagem de PL2 do informante. Nestes, com os quais se pretendeu uma análise mais fina da informação, seguiram-se, em todo o caso, os critérios que presidiram à organização dos anteriores. Assim, foram elaborados gráficos respeitantes à percentagem de palavras com erros ortográficos relativamente ao total de palavras do *Corpus* PEAPL2 (gráfico 5), à percentagem de erros com vogal posterior no total de palavras (gráfico 5.1) e à percentagem de erros com vogal posterior no total de erros ortográficos (gráfico 5.2).

No primeiro gráfico é apresentada a percentagem de palavras com todos os tipos de erro ortográfico, quer envolvam ou não vogal posterior, relativamente ao total de palavras do *Corpus* PEAPL2, obtendo-se, assim, uma noção geral da quantidade de erros num universo de 119 381 palavras.

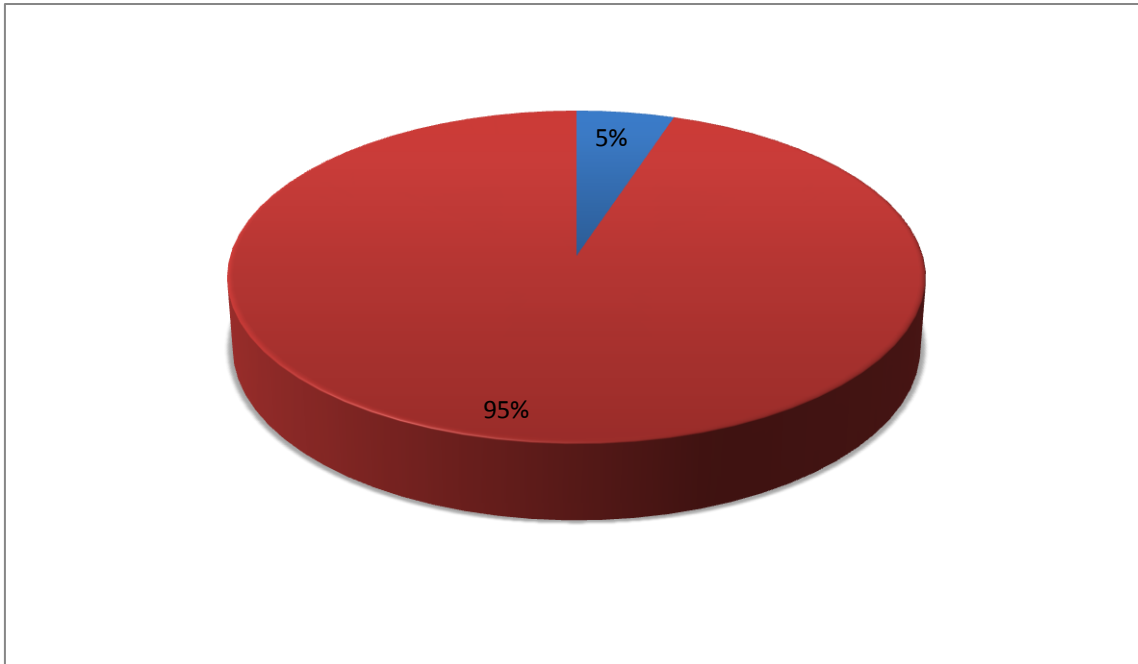


Gráfico 1: Percentagem de palavras com erros ortográficos no universo de palavras do *Corpus PEAPL2*

No gráfico 2, mais circunscrito que o anterior, está em causa um dado mais específico, a percentagem de palavras com erros envolvendo vogal posterior. Nele encontra-se calculada a percentagem deste tipo de erros no universo de palavras com erros ortográficos, ou seja, a percentagem calculada no gráfico anterior (Gráfico 1) surge agora como o universo relativamente ao qual é determinada a nova parcela.

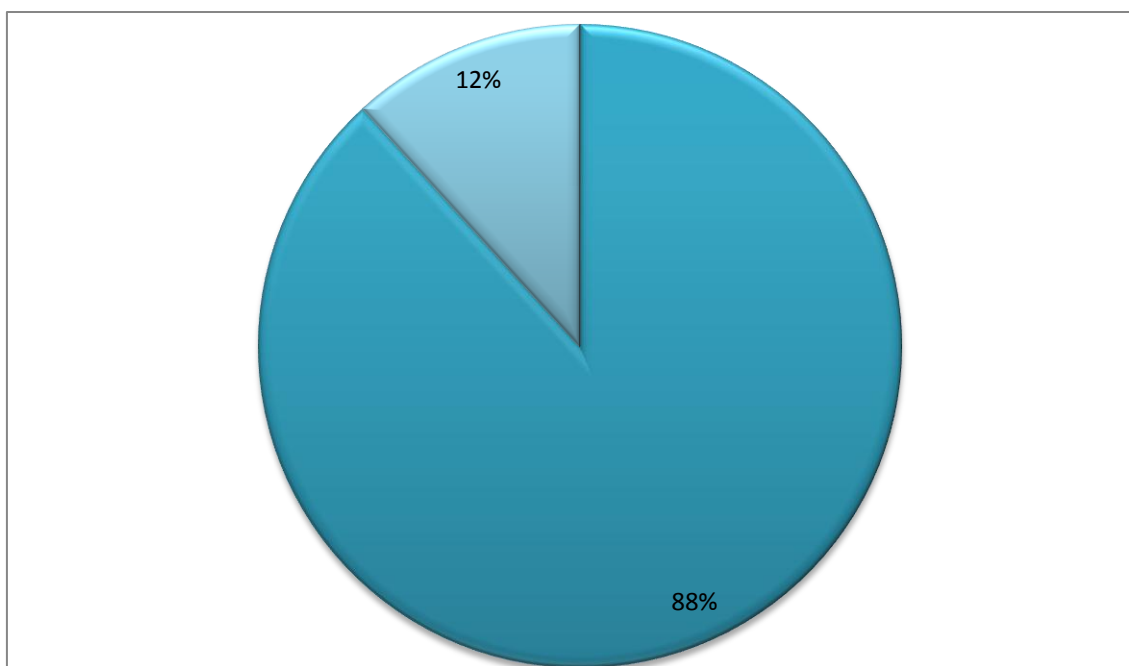


Gráfico 2: Percentagem de palavras com erros envolvendo vogal posterior relativamente ao universo de palavras com erros ortográficos

No universo de palavras dos textos produzidos pelos aprendentes de PL2, o número de palavras com erros ortográficos corresponde, então, e como se observa no gráfico 1, a uma percentagem de 5%. Este é um valor algo surpreendente, uma vez que se poderia esperar uma incidência mais elevada de desvios ortográficos neste tipo de aprendentes.

Tal situação pode ser justificada pelo facto de, neste tipo de texto, e embora haja orientação temática e restrições temporais na sua elaboração, o aluno poder seleccionar o léxico (recorrendo ao que, para si, está mais disponível). É provável que, num exercício de ditado, e apesar de essa ser uma situação em que o professor pronuncia de forma cuidada e, muitas vezes, até artificial<sup>20</sup> as palavras, os erros tivessem sido em número superior. No entanto, no momento de avaliar a correcção global dos textos, há que realçar o facto de a sua correcção ortográfica não ter impedido que, a nível sintáctico e morfológico, tivessem surgido lapsos variados, como os que se concretizam nas faltas de concordância, nominal ou verbal (aspecto cuja análise está fora de âmbito

---

<sup>20</sup> Para orientar, graficamente, os alunos, o professor opta por pronúncias que, por vezes, não coincidem com o uso real e comum.

deste trabalho). Por outro lado, há que notar que os dados apresentados se referem ao número de palavras afectadas por incorrecções ortográficas e não ao número de erros ortográficos (na verdade, algumas das palavras exibem incorrecções gráficas em mais do que um ponto da cadeia segmental).

No conjunto das palavras incorrectamente grafadas, são vários os tipos de erros detectados nas diferentes posições silábicas (ataque, núcleo ou coda) em que se encontram os elementos afectados; de entre os que envolvem consoantes, os mais representados são os que envolvem fricativas, principalmente as sibilantes, como:

- i) a substituição de <s> por <z> e vice-versa (<dezde>, <sosinha>);
- ii) o uso incorrecto de <s>, <ss> e <ç>, muitas vezes sem haver sequer coerência dentro do mesmo texto (<poça>, <esa>, <Rúsia>);
- iii) a substituição de <s> por <x> e vice-versa (<expecial>, <esperiência>);
- iv) a troca de <j> e <g> (<viajens>, <aprendizajem>);

No registo das consoantes são ainda frequentes a incorrecta representação das oclusivas: <c> por <qu> (<cuotidiano>, <cuando>).

Verificam-se também usos de (<n> pelo <m> e vice-versa quando estes grafemas representam a nasalidade da vogal anterior (<comfortavel>) e ainda, embora com frequência menor, problemas na representação de vogais de outras séries (<si>, <quasi>, <descobrendo>).

A simples análise do gráfico 1 leva-nos a concluir que os aprendentes têm um comportamento globalmente satisfatório no domínio da ortografia. Contudo, ao analisarmos o conteúdo da parcela de palavras com erros (5%), verificamos, como se pode comprovar no gráfico 2, que mais de um décimo dos erros corresponde a erros que envolvem vogal posterior. Assim, dentro do total de palavras com erros, há um grupo específico que importa para este trabalho, o das palavras que contêm erros ortográficos que envolvem a representação de vogais posteriores. No universo das palavras com erros ortográficos, a parcela de 12% corresponde, então, às que apresentam esse tipo de erros. Estes 12% englobam todos os erros com vogal posterior, independentemente do tipo de desvio e da posição acentual na qual este ocorre.

Esta parcela, que corresponde a mais de uma décima parte da totalidade dos erros, é relevante: ao tratar-se de unidades afectadas pelo processo fonológico do vocalismo átono (como já explicado anteriormente no capítulo 1), surgem condições para o erro ortográfico; por outro lado, esse processo acaba por influenciar o aprendente, que "aplica" a respectiva regra em contextos em que tal não é adequado (como se constatará ao longo deste trabalho). Por esse mesmo motivo, considera-se interessante a investigação/estudo destes desvios à norma gráfica, a fim de se tentarem estabelecer as suas causas e conseqüências, pois, como ficou atrás ilustrado, a relação entre as vogais posteriores e os grafemas que as representam não é biunívoca (as vogais fonológicas apresentam mais do que uma realização, assim como as formas gráficas envolvidas na sua representação são foneticamente polivalentes), o que leva o aprendente ao erro.

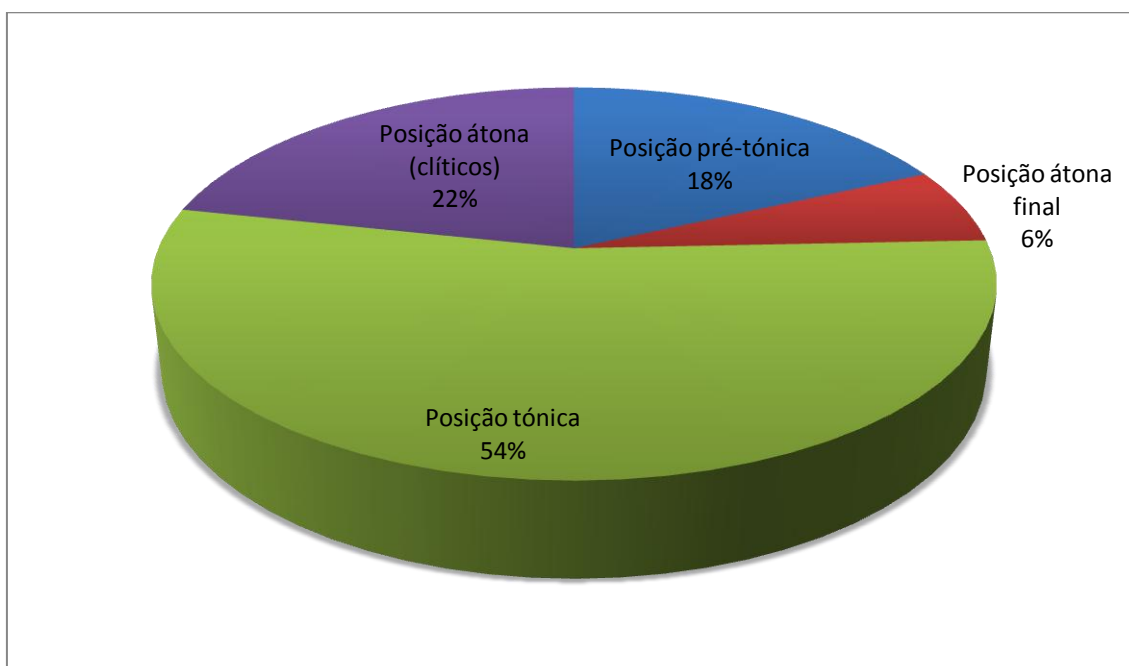


Gráfico 3: Percentagem de erros de representação segmental por posição acentual

Como já foi referido, do inventário de erros ortográficos seleccionados para o presente estudo constam erros segmentais e erros na atribuição dos sinais diacríticos; no entanto, o gráfico 3 (como o gráfico 3.1) representa unicamente os erros segmentais. Como já assinalámos, os erros acentuais estão incluídos na categoria 8, que aparece pela primeira vez representada no gráfico 4.1.

No gráfico 3, as diferentes parcelas correspondem à percentagem de erros ortográficos envolvendo a representação de vogais e semivogal da série posterior, de acordo com a respectiva posição acentual: posição tónica, posição átona final, posição pré-tónica, e, por fim, posição átona considerando unidades clíticas.

A posição acentual que reúne maior percentagem de erros é a sílaba tónica, com mais de metade dos erros registados, 54%; segue-se a posição átona relativa aos clíticos, com 22% do total dos erros. Opostamente, a posição acentual que apresenta menor número de incidências é a posição átona final, com 6% dos erros, não se manifestando sequer em todas as categorias, isto é, os erros registados nesta posição acentual apenas exemplificam as categorias 1 (<o> > <u>) e 3 (<o> > <ou>) (categorias já descritas anteriormente na secção *Procedimentos* e apresentadas na tabela 4).

Assim, o maior número de erros que envolvem vogal posterior na produção escrita recai sobre a sílaba tónica, o que não seria de esperar, pois as vogais posteriores [-alt] não sofrem, em posição acentuada, a elevação que, em posição átona, as faz convergir fonética mas não graficamente com [u] (como foi explicitado anteriormente no capítulo 1).

Note-se, no entanto, que são, principalmente, os erros segmentais registados nas formas <todo> (por <tudo>), <tudo> (por <todo>) e <foi> (por <fui>), como se pode comprovar na tabela 6, que levam a esta elevada percentagem de erros ortográficos no registo gráfico de unidades posteriores em sílaba tónica. Os dois primeiros erros estarão certamente relacionados com o facto de, em português, existirem formas variáveis dos indefinidos, que flexionam em género e número (todo(s), toda(s)) e uma forma invariável, que não flexiona, sempre pronominal (tudo), ao contrário da forma variável que pode ser determinante ou pronome. Assim sendo, estes erros sugerem um “deficit” lexical corrigível ao longo do processo de aprendizagem do aluno. Relativamente ao erro <foi> por <fui>, o facto de ambas as formas surgirem no pretérito perfeito do indicativo e de os aprendentes terem conhecimento do valor de <o>, leva-os a cometerem erros como o que está em causa.

Algo de semelhante explica os valores obtidos no registo gráfico dos clíticos: ao contrário do que seria de esperar, esta é uma parcela com uma percentagem elevada (a segunda mais elevada), devendo-se este valor, em grande parte, à frequência de um



mesmo erro: o uso de <o> em representação da conjunção alternativa <ou> representando uma redução do ditongo na variedade padrão do português europeu à qual, pelo que estes dados indiciam, os aprendentes são bastante sensíveis (como foi referido no Capítulo 1, pp. 14 e como se comprova na tabela 6).

Assim sendo, verifica-se que, nestes casos, a análise qualitativa permite perceber aspectos menos visíveis pela mera análise quantitativa. Esta situação verificar-se-á no caso de outros gráficos, motivo pelo qual se voltará a proceder a análises de tipo qualitativo.

| <b>Erros</b>      | <b>Ocorrências</b> |
|-------------------|--------------------|
| Nos (nós)         | 86                 |
| O (ou)            | 72                 |
| Todo (tudo)       | 25                 |
| Musica (música)   | 24                 |
| Foi (fui)         | 17                 |
| Tudo (todo)       | 16                 |
| Proximo (próximo) | 14                 |
| Tudos (todos)     | 12                 |
| So (só)           | 9                  |
| Publico (público) | 9                  |

Tabela 6: Dez erros ortográficos mais frequentes, envolvendo a representação de vogais posteriores, por ordem decrescente, nos textos do "*Corpus PEAPL2*"

Devido à elevada percentagem aí observada, procedeu-se seguidamente à análise dos erros registados em posição tónica, distribuindo-os pelas diferentes categorias (como se verifica no gráfico 3.1)

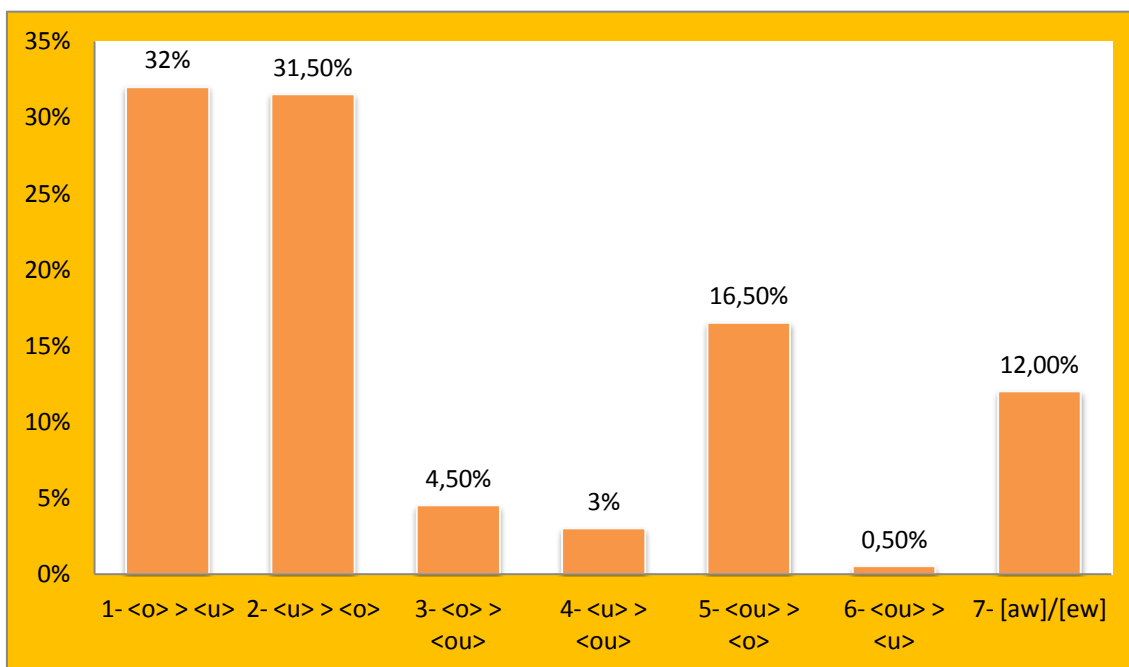


Gráfico 3.1: Percentagem de erros segmentais em posição tónica de acordo com a categoria

De forma a tentar justificar a elevada percentagem de erros em posição tónica (54%), analisou-se o conteúdo dessa parcela do gráfico 3 e elaborou-se o gráfico 3.1, no qual os erros segmentais obtidos em posição tónica surgem organizados de acordo com a categoria a que pertencem (ver tabelas 4 e 5 apresentadas anteriormente). Assim, é possível constatar se os erros se distribuem indistintamente pelas diversas categorias ou se, pelo contrário, são mais recorrentes nalguma(s). Assim sendo, o eixo vertical diz respeito à percentagem de erros dada em posição tónica e, no eixo horizontal, estão indicadas as sete categorias nas quais se incluem os referidos erros.

Deste modo, observamos que são as categorias 1 e 2, respeitantes à substituição de <o> por <u> (como em *\*muro* por *moro*) e de <u> por <o> (como em *\*dormo* por *durmo*), as que mais erros ortográficos motivam aos alunos em sílaba tónica. Contudo, estas percentagens elevadas, sobretudo nas duas primeiras categorias, são causadas pela incorrecção gráfica recorrente nalgumas palavras, como se expôs a propósito do gráfico

anterior (3). Assim, na categoria 1, os 32% de erros devem-se, principalmente, às formas <tudo> (por <todo>) e <tudos> (por <todos>), repetidas 16 e 12 vezes nos textos (como se verifica na tabela 6); já as formas <todo> (por <tudo>) e <foi> (por <fui>), que surgem 25 e 17 vezes, respectivamente, são as responsáveis pelos 31,5% de erros na categoria 2 (como se comprova na tabela 6).

Deste modo, conclui-se que a elevada percentagem de erros segmentais em posição tónica (54%) é causada principalmente por formas em que o aprendente usa <u> em vez de <o> (categoria 1) e <o> em vez de <u> (categoria 2) e também pela incorrecta representação do ditongo <ou> (representado como <o> nos textos dos aprendentes). No entanto, para os dois primeiros casos, é ainda possível observar que é um pequeno conjunto de palavras representadas frequentemente de forma incorrecta que causa as elevadas percentagens de erros (32% e 31,5%).

Importa ainda referir que a totalidade dos erros em posição tónica presente neste gráfico apenas diz respeito aos erros segmentais (em número de 200). Assim, obtém-se uma elevada percentagem de erros acentuais na posição tónica, uma vez que, no total, foram registados 554 erros na referida posição. Denota-se, deste modo, que mais de metade do total de erros registado na posição em causa diz respeito aos erros de acentuação sobre vogal tónica (354), quer seja por falta de acentuação (321) ou por uso de acentuação não prevista na grafia oficial sobre a vogal tónica (33).

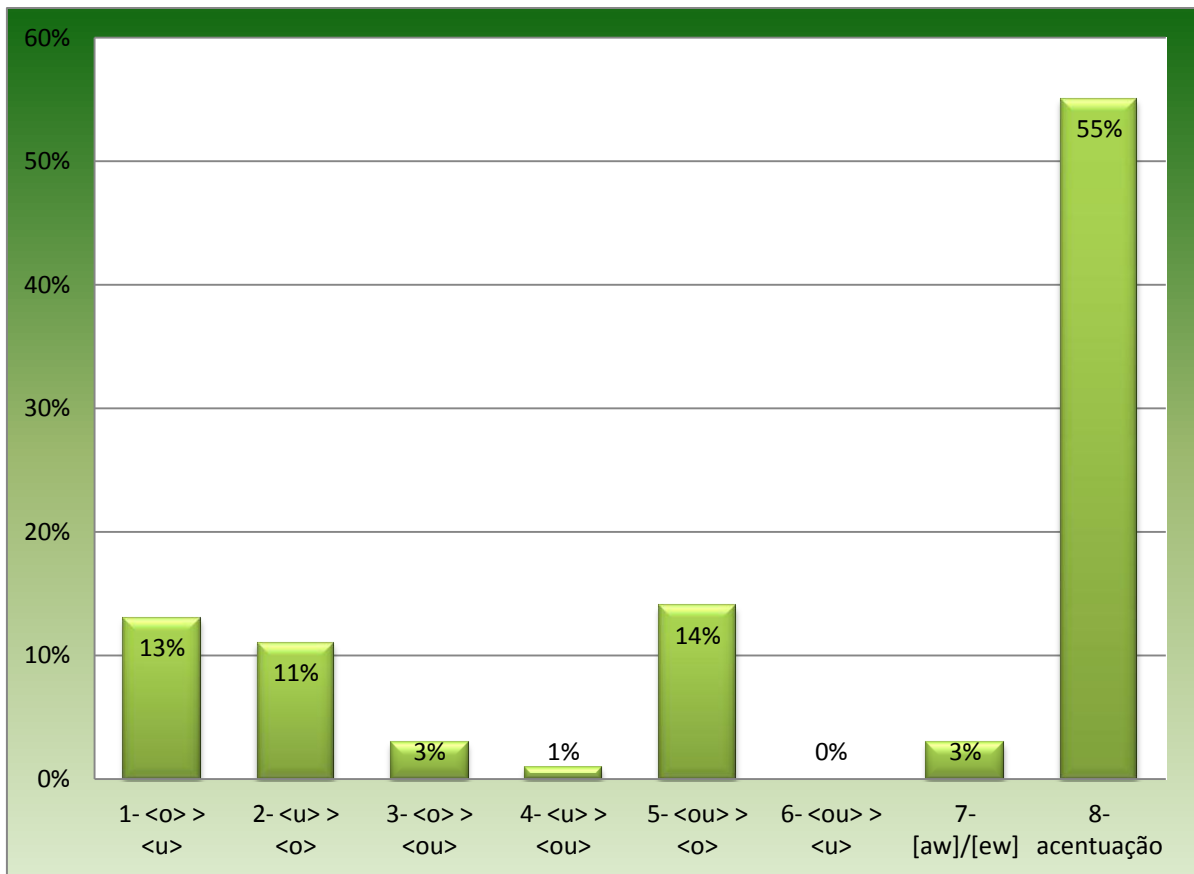


Gráfico 4: Percentagem de erros por categoria

Constatada a diversidade da natureza dos erros ortográficos que envolvem unidades posteriores, procedemos à análise dos seus valores relativos. As categorias presentes no gráfico 4 são as já anteriormente referidas (ver análise do gráfico 3.1), tendo, no presente gráfico, sido acrescentada a categoria 8, correspondente aos erros de acentuação gráfica<sup>21</sup>.

As categorias onde se verificam mais erros envolvendo vogais posteriores são as categorias 1 (<o> > <u>), com 13%, 5 (<ou> > <o>), com 14% e 8 (acentuação), com 55%. Opostamente, as categorias onde se detectam menos erros com vogais posteriores

<sup>21</sup> Recorde-se que nos gráficos 3 e 3.1 se representaram apenas os erros ortográficos de natureza segmental em posição tónica.

são as categorias 4 (1%) e 6 (0% )<sup>22</sup>, isto é, as que se referem aos desvios <u> > <ou> e <ou> > <u>, respectivamente. As categorias 3 (<ou> em vez de <o>) e 7 (representação dos ditongos [aw], [ew]) apresentam, ambas, 3% do total dos erros.

O elevado número de erros nas categorias 1 e 5 pode dever-se a inúmeros aspectos, nomeadamente, a uma grafia fonética, registando-se a pronúncia dos falantes nativos que resulta dos processos fonológicos descritos<sup>23</sup>. Assim, importa reiterar que os erros dados nestas categorias (1 e 5) denotam ser erros de ortografia fonética, em que o aprendente representa na escrita o que absorve da oralidade originando erros justificáveis, influenciados pela pronúncia (tal como foi referido anteriormente).

Por outro lado, a categoria 2 parece reflectir uma consciência da existência do erro, traduzida na incerteza dos alunos na representação gráfica. Assim, a noção de que o processo do vocalismo átono induz erros ortográficos (categoria 1) leva a um processo de hipercorreção, que conduz novamente ao erro (categoria 2). Ou seja, com o intuito de não errar, o aprendente estabelece analogias com determinados paradigmas, no caso, com o processo do vocalismo átono, acabando por substituir a forma correcta pela incorrecta: como, em posição átona, o som [u], resultante da elevação dos segmentos posteriores médios, é representado por <o>, o aprendente selecciona este grafema para registar “[u]s”, que não resultam desse processo fonológico e, portanto, devem ser grafados com <u>.

Como referido acima, a categoria 7 diz respeito à representação da semivogal [w] nos ditongos [aw] e [ew] e todos os erros ocorrem em posição tónica. Em ambos os ditongos, os alunos demonstram um conhecimento dos valores fónicos de <o>, pois denotam saber que este grafema pode representar [u] e [w]. Nesta categoria, embora a

---

<sup>22</sup> A categoria 6 é composta apenas por 2 erros ortográficos (<Lusã> em vez de <Lousã> e <muodu> em vez de <mudou>) o que perfaz uma percentagem muito baixa (0,2%) comparativamente com as restantes categorias.

<sup>23</sup> Pode, igualmente, pôr-se a hipótese de uma interferência de outras LM ou LNM's na aprendizagem da língua portuguesa ou, simplesmente, de uma incorrecta aquisição da forma fónica dos vocábulos.

percentagem de erros seja baixa, 3%, é o ditongo [ew] o que reúne maior número de erros (21).

Finalmente, a categoria 8 (relativa à acentuação) é a que reúne maior percentagem, somando 55% do total de erros com vogal posterior.

Os gráficos que se seguem (gráfico 4.1 a gráfico 4.8) dizem respeito a uma análise de pormenor dos dados já representados no gráfico 4, tendo em conta não só a diversidade do tipo de erro, como também as diferentes posições acentuais onde ele ocorre.

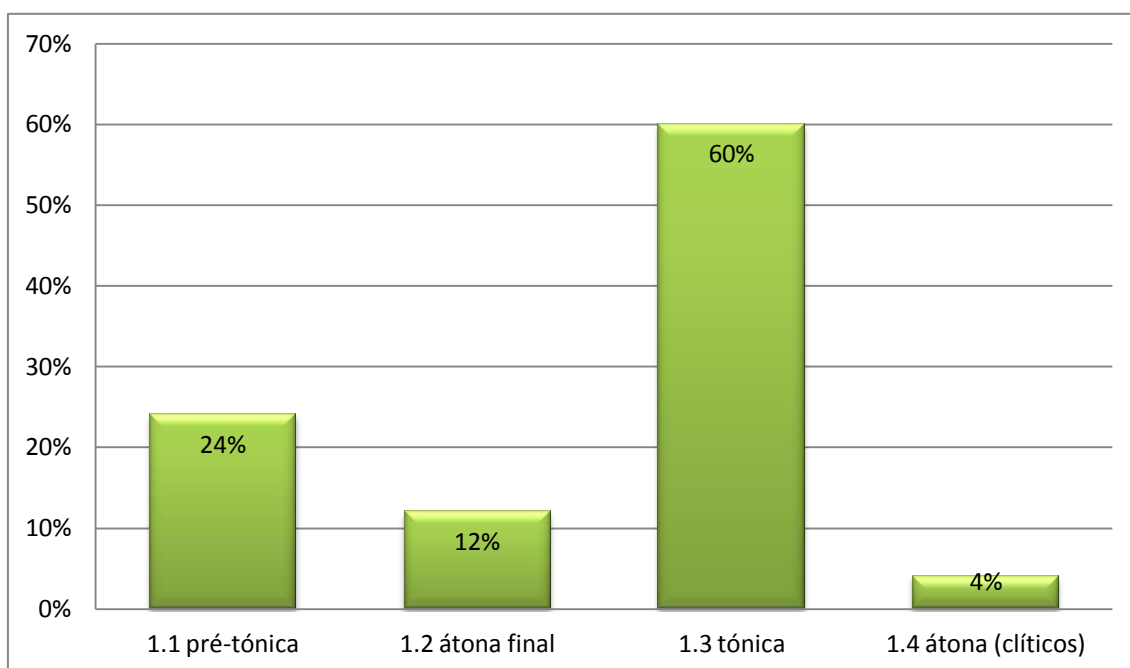


Gráfico 4.1: Percentagem de erros de categoria 1 (o > u) por posição acentual

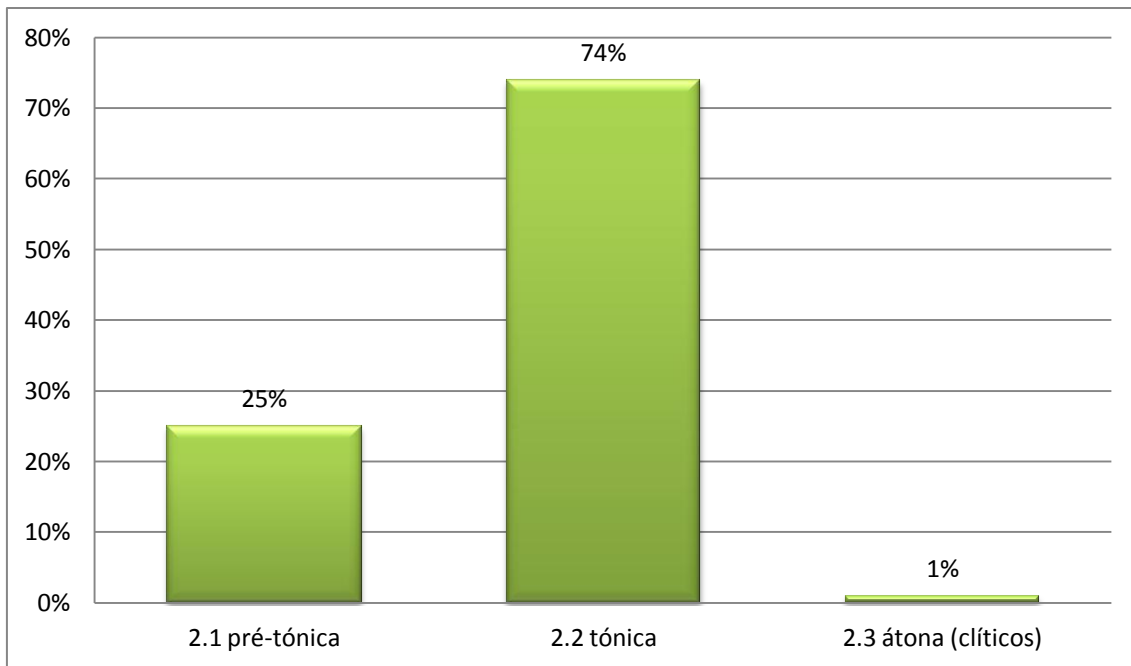


Gráfico 4.2: Percentagem de erros de categoria 2 (u > o) por posição acentual

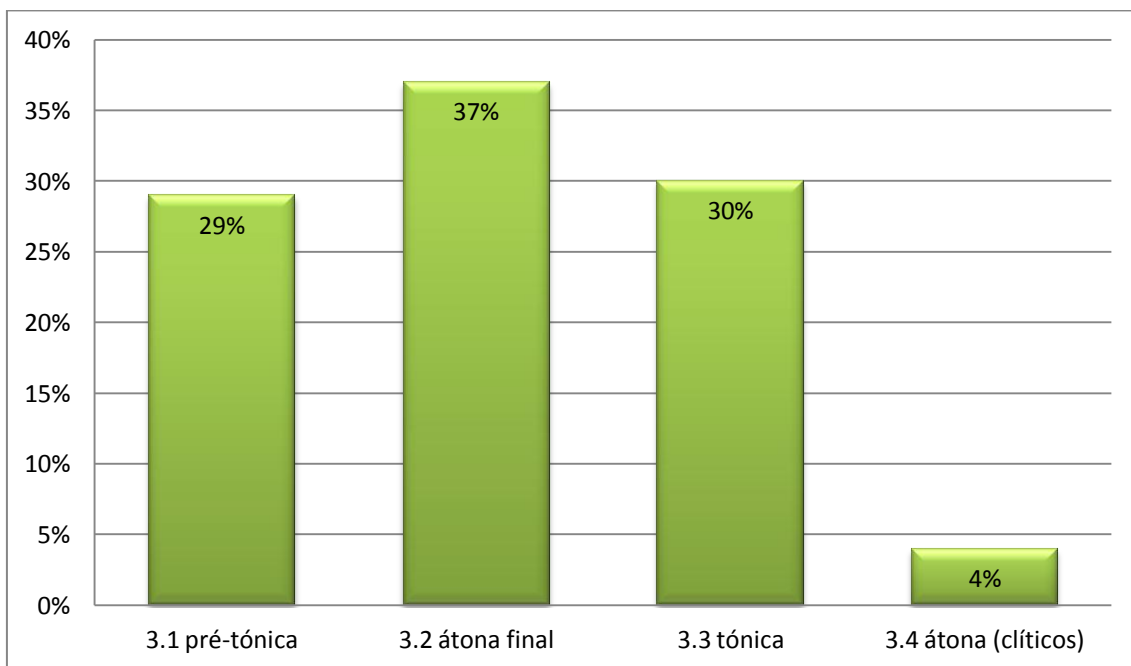


Gráfico 4.3: Percentagem de erros de categoria 3 (o > ou) por posição acentual

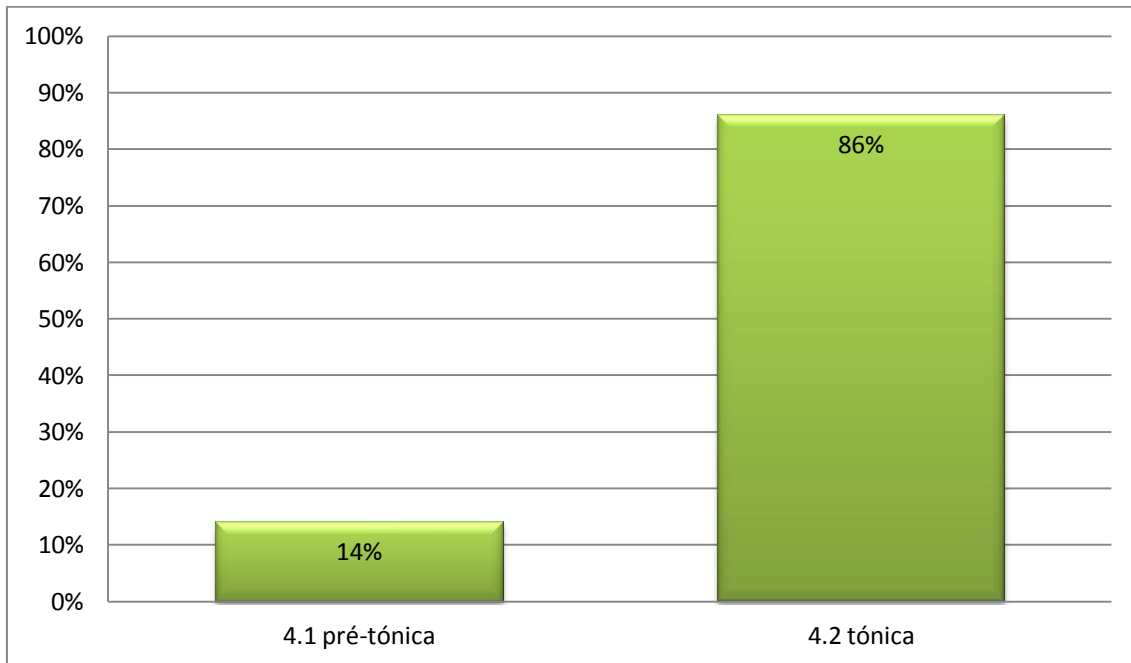


Gráfico 4.4: Percentagem de erros de categoria 4 (u > ou) por posição acentual

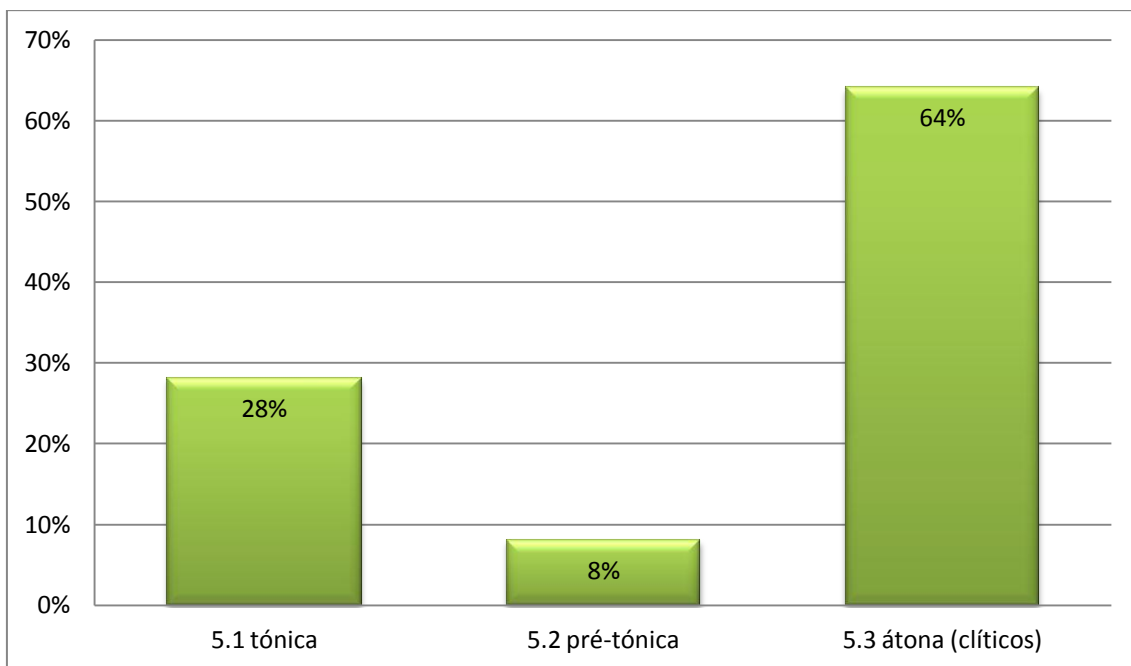


Gráfico 4.5: Percentagem de erros de categoria 5 (ou > o) por posição acentual



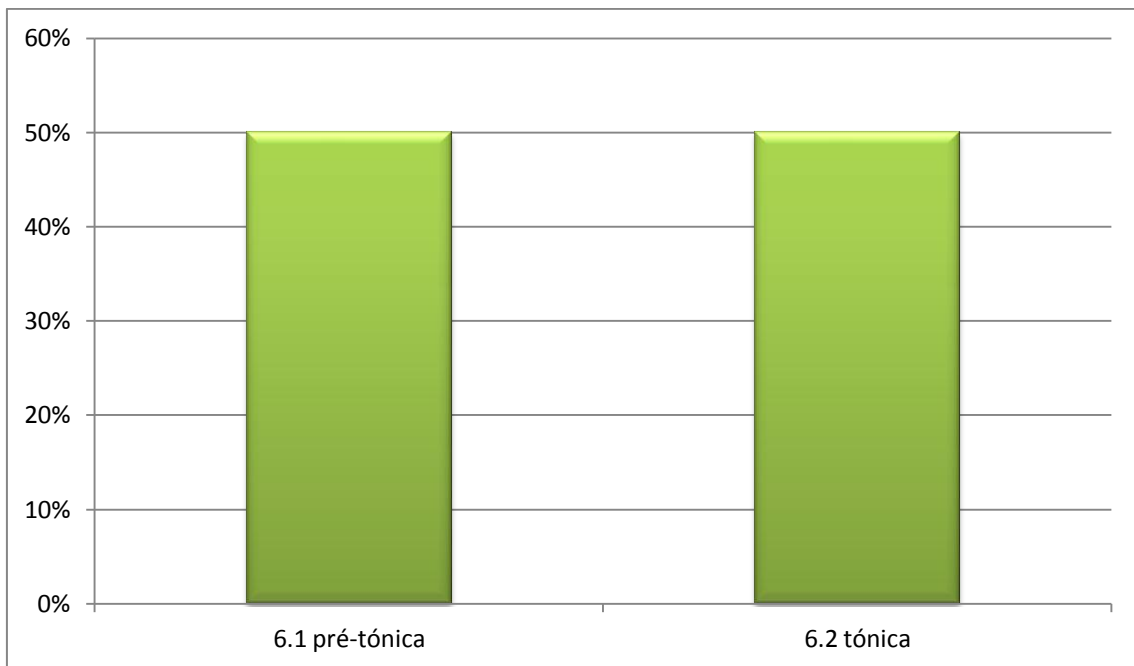


Gráfico 4.6: Percentagem de erros de categoria 6 (ou > u) por posição acentual

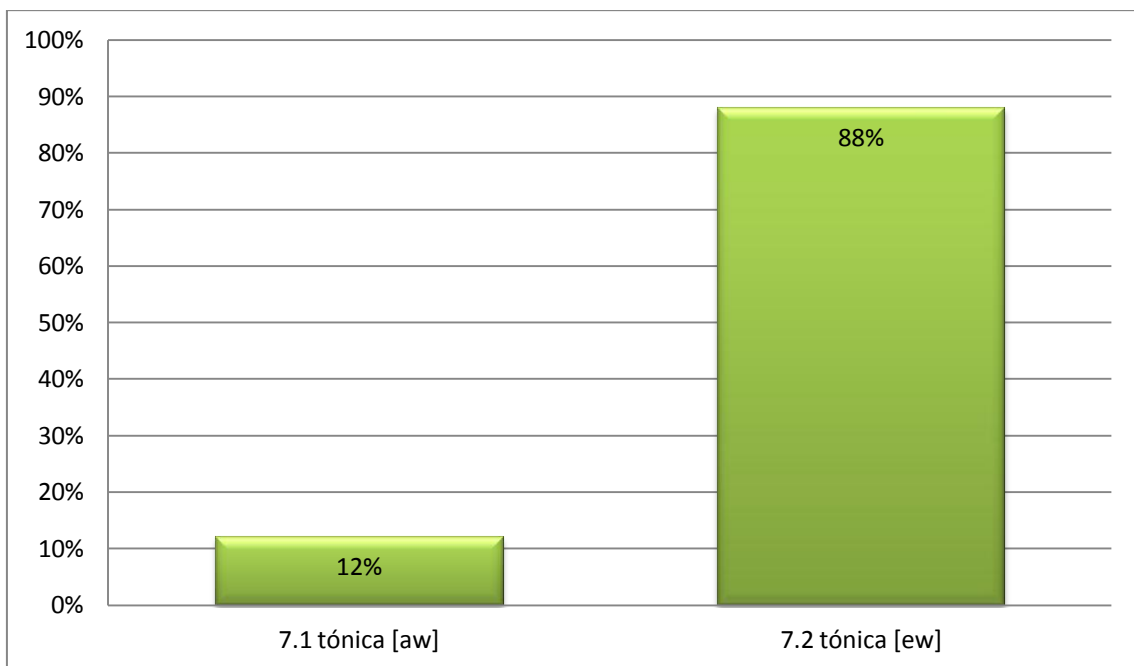


Gráfico 4.7: Percentagem de erros de categoria 7 (representações de [aw]/[ew])

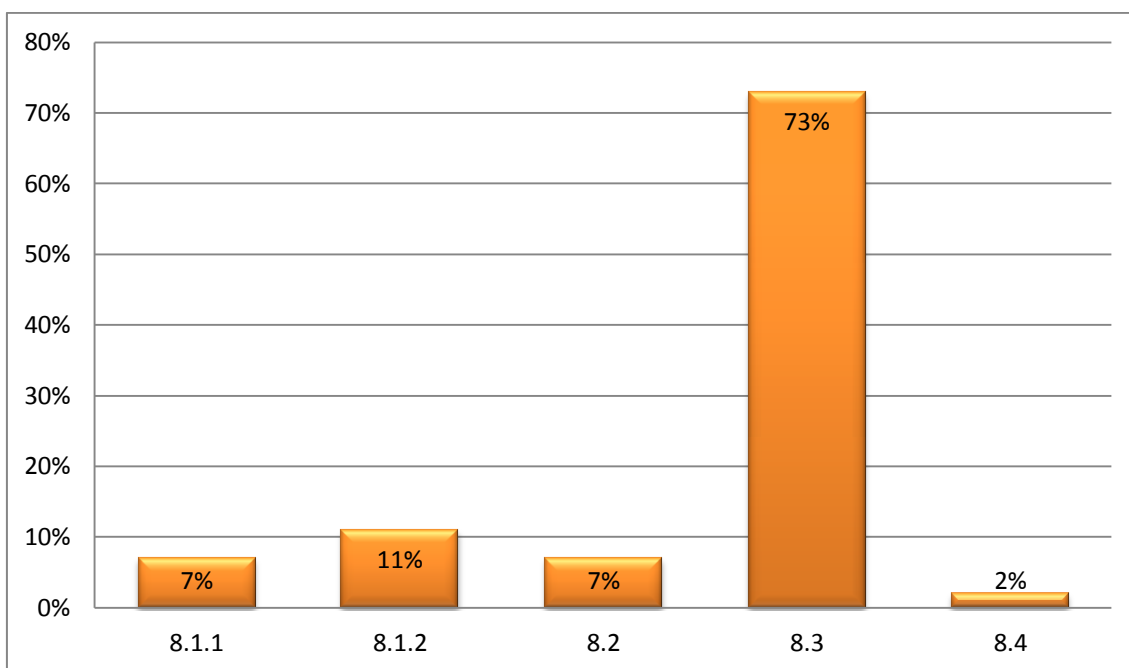


Gráfico 4.8: Percentagem de erros de categoria 8 (acentuação gráfica)

Como foi dito, devido à diversidade de erros que envolvem vogal posterior, os mesmos foram então divididos em categorias, de forma a ser possível uma análise mais minuciosa. Dentro destas categorias foi ainda possível classificá-los de acordo com a posição prosódica da sílaba em que se encontra o erro (tal como no gráfico 3). Assim, os gráficos em questão tornam possível a observação de uma categoria de erro ortográfico em função da posição acentual em que ocorre. Assim sendo, observamos que os erros de categoria 1 surgem em 4 posições acentuais, os da categoria 2, em 3 posições acentuais, os da categoria 3, em 4 posições, os da categoria 4, apenas em 2, tal como os da 6 e da 7 e os da categoria 5 se repartem por 3 posições acentuais.

Na categoria 8, embora globalmente relacionada com a acentuação gráfica, distinguiram-se, igualmente, vários tipos de erro (como foi referido anteriormente na descrição dos procedimentos e é apresentado na tabela 5).

Na categoria 1 (<o> > <u>) a maior percentagem de erros verifica-se na posição tónica (60%). Esta situação resulta de uma grafia fonética em que os aprendentes representam

graficamente as vogais tal e qual como as pronunciam, quer estas estejam em sílaba átona (\**du* em vez de *do*) ou em sílaba tónica (\**durmir* em vez de *dormir*, \**muro* em vez de *moro*), pois, inúmeras vezes, ao terem noção dos valores de <o>, e tentando corrigir-se, incorrem no erro, mesmo sendo a sílaba forte da palavra (\**fugem* em vez de *fogem*, \**gusto* em vez de *gosto*, \**punto* em vez de *ponto*). O mesmo acontece na categoria 2, em que é a posição tónica a que acarreta maior número de erros e com uma elevada percentagem, 74%.

Na categoria 3, é a sílaba átona final aquela em que se manifestam mais erros (37%); note-se que esta é a única categoria onde esta posição acentual se destaca, contrariando a tendência para a pouca expressividade dos erros ortográficos em causa nessa posição, pois neste tipo de sílabas não se registam erros de todas as categorias. Verifica-se também que os erros da categoria 3, nesta posição acentual, dizem todos respeito a verbos involuntariamente apresentados no pretérito perfeito, quando a forma correcta seria a do presente (\**passou* em vez de *passo*, \**gostou* em vez de *gosto*). Esta elevada percentagem pode, uma vez mais, ser fruto de alguma insegurança relativamente aos valores de <ou>.

Na categoria 4, a percentagem mais alta de erros regista-se na sílaba tónica (86%) e na categoria 5, na átona relativa aos clíticos (64%) (também esta incidência nesta posição acentual representa uma situação de excepção, pois através da presente análise constata-se que é a sílaba tónica a que reúne maior número de erros, sendo que aquela percentagem elevada se deve, unicamente, a um erro, <o> equivalente a <ou>, conjunção, bastante frequente nos textos do *Corpus*).

Na categoria 6, verifica-se uma equidade de percentagem de erros nas diferentes posições: 50%, quer em posição pré-tónica, quer em posição tónica; na categoria 7, dedicada à representação dos ditongos [aw] e [ew] (<au> e <eu>, respectivamente), destaca-se uma maior percentagem de erros envolvendo o ditongo [ew] (88%).

Assim, há posições acentuais onde não se registam algumas categorias de erros, como é o caso da posição átona final ou dos clíticos. A categoria 6 é a menos representada (0,2%), isto é, é aquela que regista menor quantidade de erros envolvendo vogal posterior e, conseqüentemente, menos diversidade nas posições acentuais onde eles ocorrem. Situação semelhante se verifica com a categoria 4 (situação em que o

aprendente em vez de <u> utilizou, incorrectamente, o ditongo <ou>) nas posições pré-tónica e tónica.

De uma forma geral, como já dissemos, a maior incidência dos erros dá-se na sílaba tónica, excepto na categoria 5, em que a sílaba que se destaca por registar maior percentagem de erros é a átona (clíticos) e na categoria 3 (o > ou), em que é a sílaba átona final a que regista mais erros. Depois da sílaba tónica, segue-se a posição pré-tónica como posição acentual com mais ocorrência de erros nas diversas categorias. Deste modo, a posição pré-tónica demonstra ser a posição átona com mais percentagem de erros, o que poderá dever-se a uma forte influência do processo do vocalismo átono.

No que concerne à percentagem de erros em sílaba tónica, constata-se que, além de aí ocorrerem erros de todas as categorias, essa posição é também aquela que apresenta as percentagens de erros mais elevadas no cenário geral, situação algo inesperada, como foi referido anteriormente. Igualmente surpreendente é a elevada (a mais alta aliás) percentagem de erros da categoria 3 registada em sílaba átona final onde apenas ocorrem erros duas categorias, 1 (<o> > <u>) e 3 (<o> > <ou>) e, como se observa no gráfico 3, as valores globais dos erros são os mais baixos dos obtidos nas diferentes posições acentuais. Quanto aos erros em sílaba átona (clíticos), a manifestação destes nas diferentes categorias é baixa, devido ao facto de os clíticos serem em número reduzido em português e usados muito frequentemente, o que facilita e acelera a memorização das respectivas formas gráficas (4% nas categorias 1 (<o> > <u>) e 3 (<o> > <ou>), 1% na categoria 2 (<u> > <o>)). A excepção, a que já nos referimos atrás, é feita para a categoria 5 (<ou> > <o>), com 64% dos erros.

Assim, a posição átona (clíticos) é a que demonstra percentagens mais baixas de erros ortográficos; os aprendentes denotam não ter muitas dúvidas em vocábulos com este estatuto acentual (e gramatical), assim como nas sílabas em posição pós-tónica. Pelo contrário, manifestam elevadas percentagens de erros na sílaba tónica, contrariando-se assim as expectativas (como já foi referido anteriormente na análise do gráfico 3).

Relativamente à acentuação, o uso do tipo incorrecto de diacrítico é o erro menos frequente dentro da categoria dos erros acentuais (2%), por oposição à falta de acento gráfico, que é o mais frequente (73%); embora todas as parcelas/subcategorias digam respeito a erros envolvendo o diacrítico, importa referir o facto de o primeiro (8.1.1) e o

terceiro (8.2) tipo de desvios, apesar de serem de natureza diferente, pois uns recaem sobre a vogal átona (erros causados por acento gráfico ou til sobre vogal ou semivogal átonas por deslocação do acento ou til em palavras que o contenham) e outros sobre a vogal tónica (erros causados por atribuição de acento gráfico ou til sobre vogal tónica em palavras em que esse acento não está previsto), terem a mesma percentagem de erros, 7%.

Os gráficos (5, 5.1 e 5.2) que a seguir apresentamos e analisamos mostram a percentagem de (palavras com) erros dados pelos aprendentes de acordo com o seu nível de aprendizagem. Neles é possível analisar, em diferentes cenários, que vão de um universo mais abrangente para um mais particular, a quantidade de formas gráficas desviantes desde o início do processo de aprendizagem até um nível mais avançado.

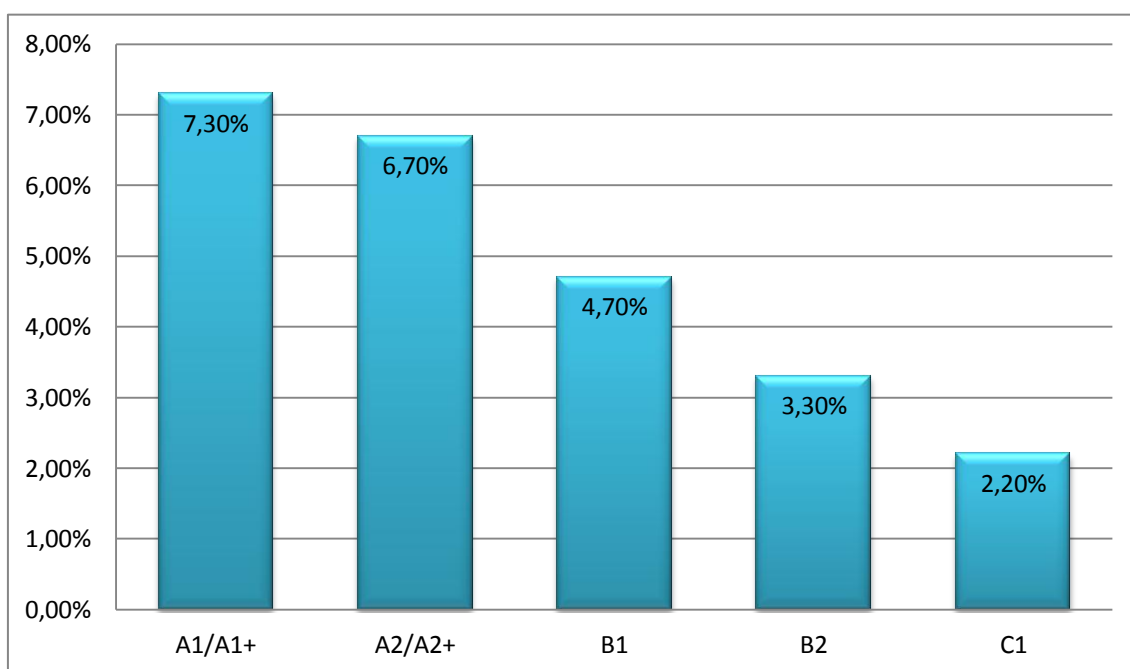


Gráfico 5: Percentagem de palavras com erros ortográficos no universo de palavras obtidas por nível de aprendizagem dos informantes

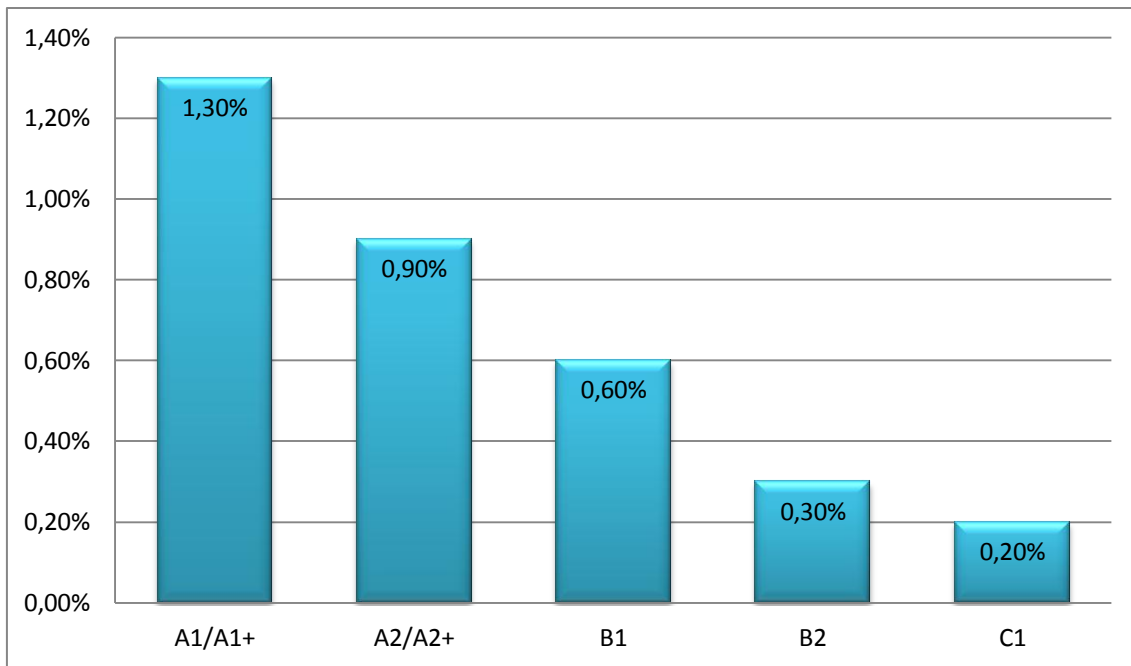


Gráfico 5.1: Percentagem de erros ortográficos envolvendo vogal posterior no universo de palavras obtidas por nível de aprendizagem dos informantes

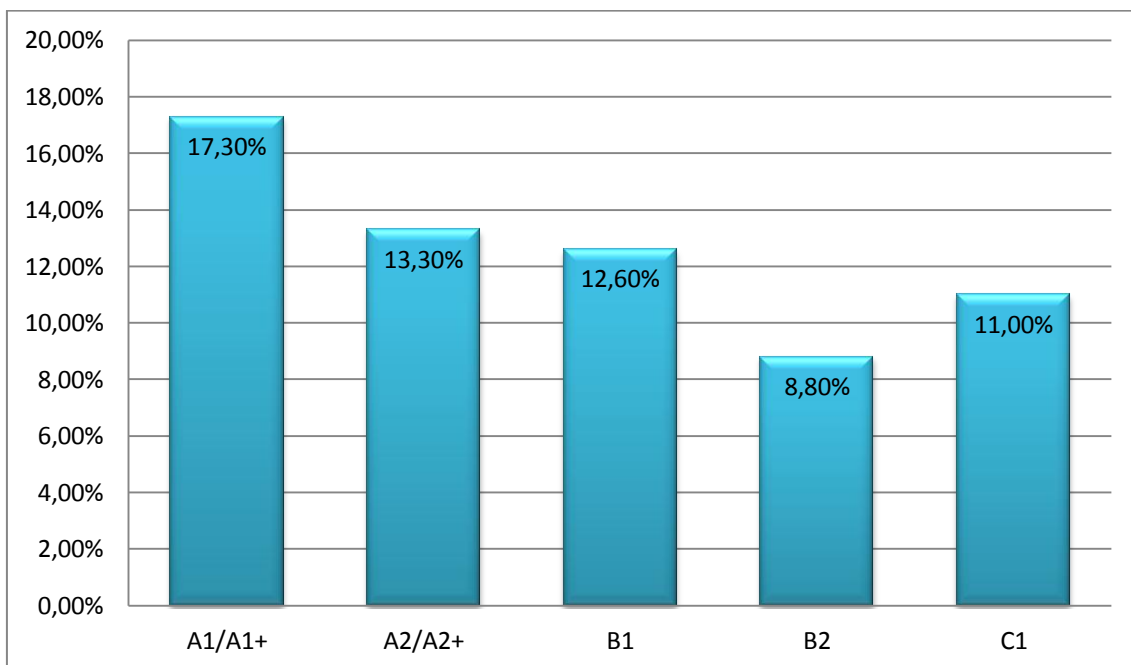


Gráfico 5.2: Percentagem de erros ortográficos envolvendo vogal posterior no universo de palavras com erros ortográficos por nível de aprendizagem dos informantes

Os três gráficos representam percentagens de (palavras com) erros dados por aprendentes de PL2, considerando os diversos níveis de aprendizagem, de A1 a C1. No que concerne ao gráfico 5, este é relativo à percentagem de palavras com erros ortográficos, de acordo com o nível de aprendizagem, no universo correspondente à totalidade das palavras dos textos produzidos pelos alunos desse mesmo nível; o gráfico 5.1 diz respeito à percentagem de erros ortográficos envolvendo vogal posterior por nível de aprendizagem, no mesmo universo; finalmente o gráfico 5.2 representa a percentagem de erros ortográficos envolvendo vogal posterior por nível de aprendizagem mas, neste caso, no universo de palavras com erros ortográficos produzidos por cada nível. Assim sendo, os gráficos 5 e 5.1 têm em comum o universo a partir do qual se calcula a percentagem de erros, embora no primeiro (5) estejam em causa as palavras com qualquer tipo de erro ortográfico e, no segundo (5.1), somente os erros com vogal posterior; no gráfico 5.2 o universo difere dos anteriores, já que a percentagem de erros com vogal posterior por nível é aí calculada no universo do total de erros ortográficos produzidos por nível de aprendizagem. Embora semelhantes, estes gráficos representam diferentes organizações dos dados, pois enquanto no primeiro (5), o mais geral, se vê a percentagem de todos os tipos de erros, no segundo (5.1), já está em causa um grupo de erros em concreto, os erros que envolvem vogal posterior, no universo das palavras produzidas e, no último gráfico (5.2), o mais concreto/específico, estão representados novamente os erros com vogal posterior mas, desta feita, num panorama particular, o do universo de palavras com erros.

Assim, nos gráficos 5 e 5.1 verifica-se uma tendência decrescente na percentagem de (palavras com) erros, inferindo-se uma evolução e dissipação das dúvidas ao longo do processo de aprendizagem. Tanto nos erros ortográficos em geral, como nos erros que envolvem vogal posterior, se assinala uma progressão; à medida que desenvolvem os seus conhecimentos da língua e se intensifica a sua exposição a *input* escrito, os aprendentes corrigem/aperfeiçoam a sua ortografia.

A tendência mostrada nestes gráficos (5 e 5.1) é uma tendência geral e previsível. À medida que os informantes evoluem no nível de aprendizagem, a quantidade de erros ortográficos que cometem diminui também, supondo-se, portanto, que caso houvesse

aprendentes correspondentes ao nível de proficiência C2, a percentagem de erros seria inferior à que se regista no nível C1 (2,2%).

Importa ainda realçar o facto de, no gráfico 5, as percentagens irem dos 7,3% aos 2,2%, enquanto no gráfico 5.1 as percentagens se apresentam naturalmente mais baixas, uma vez que o primeiro gráfico é referente a todos os tipos de erros, quer envolvam ou não vogal posterior, havendo, por isso, há mais erros por nível, enquanto o segundo gráfico diz somente respeito aos erros que envolvem vogal posterior, ou seja, a uma parte do conjunto global dos erros; os valores oscilam, então, entre 1,3% e 0,2%.

No gráfico 5.2, relativo aos erros com vogal posterior no universo dos erros ortográficos, esta tendência é contrariada. Embora, globalmente, se registre a diminuição da percentagem de erros ao longo do processo de aprendizagem, também se detecta uma pequena inversão na passagem do nível B2 (8,8%) para o nível C1 (11%); quando os dados pareciam indicar um *continuum* evolutivo regista-se, no nível mais avançado deste *Corpus*, um retrocesso. Assim, no nível onde se esperaria um menor número de erros, surpreendentemente a percentagem não só não diminui como é mais elevada que no nível anterior.

Uma análise dos intervalos entre os valores de erros registados, no gráfico 5.2, para os informantes de cada um dos níveis de aprendizagem, permite, ainda, observar que estes são maiores entre A1/A1+ e A2/A2+ (17,3% *versus* 13,3%) e entre B1 e B2 (12,6% *versus* 8,8%) do que entre A2/A2+ e B1 (13,3% *versus* 12,6%).

Assim, os dados registados nos gráficos 5 e 5.1 reflectem a tendência para um crescente domínio do sistema ortográfico em geral, enquanto no gráfico 5.2 está patente um ligeiro e aparente retrocesso do nível B2 para o nível C1. Contudo, neste último nível de aprendizagem, os erros encontrados dizem maioritariamente respeito a erros acentuais (12 erros), sendo que os erros segmentais são uma quantidade mínima (4 erros). Assim sendo, não se tratará de um verdadeiro retrocesso no que diz respeito ao domínio do aspecto da ortografia do português que está em apreciação neste trabalho, já que as dúvidas dos aprendentes recaem essencialmente sobre a acentuação.



Assim, após a análise dos dados, é possível observar que, em termos de grau de proficiência em PL2, são os alunos do nível A1 aqueles que dão o maior número de erros, verificando-se, regra geral, uma diminuição desse número ao longo do processo de aprendizagem. Depreende-se, também, que a maior parte dos erros segmentais se incluem nas categorias 1 (<o> > <u>) e 5 (<ou> > <o>) manifestando-se assim dificuldades por parte dos aprendentes quer na representação da vogal <o>, quer na representação do ditongo <ou>, como foi referido ao longo da análise.. Ainda relativamente aos mesmos erros (segmentais), verifica-se que a maioria destes recai na posição tónica, ao contrário do que seria de esperar, pois seria mais provável as sílabas átonas possuírem a maioria dos erros devido às alterações fonéticas provocadas pelo processo do vocalismo átono. Assim sendo, em alguns casos, as anomalias ortográficas podem reflectir uma pronúncia informal e por isso descuidada, onde os fenómenos de dissimilação e de assimilação<sup>24</sup>, entre outros, abundam, motivando uma escrita fonética. Por fim, observa-se ainda que os erros acentuais têm também uma forte representatividade no conjunto dos erros ortográficos analisados.

Através da observação dos dados apresentados no anexo 2, é possível concluir-se, igualmente, que são os aprendentes cujas línguas maternas são línguas românicas, principalmente italianos e espanhóis, aqueles que dão maior número de erros ortográficos nos textos que constituem o *Corpus*. Este facto deve-se eventualmente à proximidade e similitude do italiano e do espanhol ao português. Situação inversa se verifica relativamente a alguns aprendentes oriundos do continente asiático, devido às diferenças abissais das línguas e dos sistemas de escrita, nomeadamente pelo facto de o chinês e o japonês terem sistemas de escrita não fonológicos, isto é, sistemas a cujos grafemas correspondem unidades do plano morfológico e / ou lexical e não unidades do plano fónico; já o português possui um sistema alfabético de escrita, onde as letras correspondem a unidades fónicas. Tais dissemelhanças dificultam qualquer influência, estrutural ou ortográfica, sobre estes aprendentes, que não se podem guiar por analogias

---

<sup>24</sup> Entende-se por dissimilação a diferenciação de um som por influência de um som vizinho, articulatoriamente próximo, com o objectivo de se distinguir dele procurando-se, assim, evitar uma sequência de dois fonemas semelhantes dentro da mesma palavra. A assimilação, pelo contrário, consiste na modificação de um som por influência do som vizinho que com ele passa a partilhar traços articulatorios, tornando-se foneticamente parecidos (Mateus, 2005: 230).

ou parecenças de uma língua com a outra. Assim, aquelas semelhanças que levam muitos aprendentes ao erro não afectam aprendentes orientais, sendo estes os alunos que manifestam menor percentagem de erros ortográficos.

## Conclusão

Após a análise do *Corpus* PEAPL2 e o apuramento dos textos relevantes em função dos objectivos desta dissertação, é possível concluir-se neste trabalho, essencialmente empírico, que, dentro do universo dos erros dados pelos aprendentes de PLN, os erros envolvendo vogais posteriores são percentualmente significativos, distribuindo-se entre segmentais e acentuais. Conclui-se, igualmente, que são os erros acentuais, e os erros em que o aprendente usa as formas gráficas <u>, em vez de <o>, e <o>, em vez do ditongo <ou>, resultado da adopção de uma grafia fonética, os que reúnem maior percentagem, bem como é a posição tónica aquela que agrega maior número de desvios ortográficos; nas posições átonas, são os clíticos que registam a maior percentagem de erros. Por outro lado, dentro da categoria dos erros acentuais, a falta de acento gráfico é o erro mais comum.

Constatou-se, também, que são os aprendentes dos níveis iniciais de aprendizagem do português, ou seja, os níveis A1, A2 e A2+, os que dão maior número de erros. Deste modo, segue-se a tendência previsível verificando-se uma evolução ao longo do processo de aprendizagem, na qual o aprendente se torna cada vez mais proficiente, traduzindo-se esta progressão não só na diminuição da quantidade de erros, como nas estruturas sintácticas utilizadas.

Assim, após a análise minuciosa dos gráficos, dos tipos de erro e da sua frequência, pode colocar-se a hipótese da influência dos sistemas fonológicos e gráficos das respectivas Línguas Maternas (LM) dos aprendentes que, por um lado, pode ser um factor de facilitação devido às semelhanças na forma gráfica, e, por outro lado, pode afigurar-se um obstáculo que entrava a aprendizagem do português enquanto língua não materna (PLN), pois reflecte-se erradamente contribuindo para uma maior quantidade de erros. Esta hipótese, não tendo sido investigada neste trabalho, constitui-se, ainda assim, como uma possibilidade de desenvolvimento futuro nesta área de estudos.

No entanto, não só as respectivas LM dos aprendentes parecem interferir, por vezes negativamente, na escrita. Também a oralidade, quer a dos falantes nativos do

português, quer ainda a dos próprios aprendentes pode servir de modelo a uma grafia de base fonética.

Assim, para a aprendizagem da ortografia do português é preciso ter presente a sua base fonológica e dominar igualmente os mecanismos que, no plano da actualização fonética, conduzem a alguns desajustes. Melhorar o conhecimento da estrutura morfológica e lexical da língua facilitará, igualmente, a aprendizagem desta competência.

Exercendo, então, a oralidade tão grande influência sobre a escrita, note-se que a maioria dos erros identificados neste estudo são erros fónicos (Rio-Torto, 1998: 607-610) causados por uma reprodução do fonetismo da palavra ou causados por hipercorreção. Assim sendo, a produção da linguagem oral, principalmente, por parte dos docentes, deve ser bastante cuidada, clara e rigorosa de modo a poder orientar adequadamente o aprendente no seu registo gráfico.

Tendo este trabalho tido como objectivo fulcral a organização dos dados atrás descritos de forma a ilustrar que a questão da aprendizagem da ortografia que merece atenção por parte dos investigadores, há, de facto, várias hipóteses de continuidade. Assim, e como já foi referido, estudos futuros poderão dar seguimento à análise da relação dos erros com as respectivas LM, assunto não explorado neste trabalho. Outra hipótese será investigar a classe de palavras sobre que incidem os erros (substantivos, pronomes, verbos, entre outras), de acordo com o nível de aprendizagem. Uma outra área de trabalho seria ainda pesquisar-se a relação entre o número de erros dados pelos aprendentes de PLN e o seu tempo de aprendizagem do português ou a frequência com que o utilizam.

## Bibliografia

BARBOSA, Jorge Morais (1994) *Introdução ao Estudo da Fonologia e Morfologia do Português*. Coimbra: Livraria Almedina.

CÂMARA Jr., Joaquim Mattoso (1996) *Estrutura da Língua Portuguesa*, 25<sup>a</sup> ed.. Petrópolis, RJ: Editora Vozes.

CHOMSKY, Noam, HALLE, Morris (1968) *The Sound Pattern of English*. London: Harper & Row.

Conselho da Europa (2001) *Quadro Europeu Comum de Referência para as Línguas*. Aprendizagem, ensino, avaliação. Lisboa: Asa.

LEIRIA, Isabel (2004) *Português língua segunda e língua estrangeira: investigação e ensino*. In: *Idiomático. Revista Digital de Didáctica de PLNM*, nº 3. Centro Virtual Camões. Disponível em <http://www.instituto-camoes.pt/cvc/idiomatico/03/portuguesLSeLE.pdf>.

MADDIESON Ian (2009) *Patterns of Sounds*. Cambridge: Cambridge University Press.

MARQUILHAS, Rita (2003) *Mudança Analógica e Elevação das Vogais pretónicas*. In org. de Ivo Castro e Inês Duarte, *Razões e Emoção, miscelânea de estudos em*

*homenagem a Maria Helena Mira Mateus*, vol. II. Lisboa: Imprensa Nacional Casa da Moeda.

MARTINS, Cristina (coord.) (2010) *Corpus de Produções Escritas de Português L2 (PEAPL2)*. Disponível, a 10.09.2011, em <http://www.uc.pt/fluc/rcpl2/>.

MARTINS, Cristina (2011, no prelo) *O Corpus de Produções Escritas de Aprendentes de PL2 (PEAPL2/CELGA). Caracterização e desenvolvimento de uma infra-estrutura de investigação*. In: FLORES, Cristina (ed.), *Português Língua Não Materna: Investigação e ensino*. Lisboa: Lidel.

MATEUS, Maria Helena Mira (1996) *Fonologia*. In Isabel Hub Faria et al., *Introdução à Linguística Geral e Portuguesa*. Lisboa: Caminho.

MATEUS, Maria Helena Mira (2002) *A face exposta da língua portuguesa*. Lisboa: Imprensa Nacional - Casa da Moeda.

MATEUS, Maria Helena M. (2006a) *Fonologia*. In Maria Helena Mira Mateus et al., *Gramática da Língua Portuguesa*. 7ª ed.. Lisboa: Editorial Caminho.

MATEUS, Maria Helena Mira (2006b) *Sobre a natureza fonológica da ortografia portuguesa*. In *Estudos da linguagem: questões de fonética e fonologia. Uma homenagem a Luís Carlos Cagliari*. Bahia: Universidade Estadual do Sudoeste da Bahia.

MATEUS, Maria Helena, *A Pronúncia do Português Europeu* (2005) disponível a 19.09.2011 informação em: [http://cvc.instituto-camoes.pt/cpp2/acessibilidade/capitulo3\\_2.html](http://cvc.instituto-camoes.pt/cpp2/acessibilidade/capitulo3_2.html) e [http://cvc.instituto-camoes.pt/cpp2/acessibilidade/capitulo9\\_1.html](http://cvc.instituto-camoes.pt/cpp2/acessibilidade/capitulo9_1.html).

MATEUS, Maria Helena, D'ANDRADE, Ernesto (2000) *The phonology of portuguese*. Oxford: Oxford University Press.

MATEUS, Maria Helena M. *et al.* (2005) *Fonética e Fonologia do Português*. Lisboa: Universidade Aberta.

RIO-TORTO, Graça Maria (2000) *Para uma pedagogia do erro*. In *Didáctica da Língua e da Literatura, vol I. Actas do V Congresso Internacional de Didáctica da Língua e da Literatura*. Coimbra: Livraria Almedina/ILLP-FLUC, pp. 595-618.