Prosodic phrasing across varieties of European Portuguese: segmental evidence

Marina Vigário, Marisa Cruz, Pedro Oliveira, Nuno Paulino
Universidade de Lisboa

InAPoP Project (PTDC/CLE-LIN/119787/2010)
Outline

■ Background
  - prosodic segmental phenomena across E. Portuguese (EP) dialects
  - prosodic domains and rules in EP
  - Glide insertion to break a hiatus
    and word final epenthesis

■ Method: materials and data collection

■ Results and discussion

■ Main Conclusions
Prosodic phenomena across E. Portuguese dialects

Standard EP
- Syllable/morpheme final fricatives resyllabification
  \[ BE\text{la}s \text{ Obras} > BE\text{la}[z] \text{ Obras} \]
- Word-final vowel deletion
  \[ GA\text{to} a\text{MIgo} > GA\text{t aMIgo} \]
- High V semivocalization (V V\text{>G} V)
  \[ GA\text{to aMIgo} > Ga\text{t}[w] a\text{MIgo} \]
- Syllable degemination
  \[ CA\text{Mpo poluÍdo} > CA\text{MpoluÍdo} \]

(Sá Nogueira 1945, 1948; Andrade & Viana 1994; Frota 2000; Mateus & Andrade 2000)
Prosodic phenomena across E. Portuguese dialects

- Sandhi/prosodic processes are not among the typical phenomena observed in dialectal studies more tuned to segmental, lexical, and also to a certain extent morphological and syntactic variation (Boléo & Silva 1962; Cintra 1971; Rodrigues 2003; Saramago 2006; Aguiar 2009; Carrilho et al. 2010)

- confined to observations disperse in the literature
  \( \rightarrow \) the precise phonological conditions for sandhi phenomena often unknown;
  \( \rightarrow \) exact geographical limits largely to be determined
Prosodic phenomena across E. Portuguese dialects

Prosodic phenomena in non-Standard dialects of Portuguese

Word final fricative [ʒ] (instead of [z]) when followed by word starting in a vowel

\[ \text{as aulas} \rightarrow \text{a [ʒ] aulas} \]

North (e.g. Trás-os-Montes - Aguiar 2009)

Glide insertion to break a hiatus (a_á)

\[ \text{a Aula} \rightarrow \text{a [j] Aula} \]

North (e.g. Vila Real Trás-os-Montes, Beira Interior Norte – Lopo 1895, Pereira 1908; Santos 1897)

[i] or [ɨ] insertion after oxitone word

\[ \text{café} \rightarrow \text{café[i]} \]

Alentejo (Vasconcelos 1896, 1987; Paiva Boléo & M.H. Santos Silva 1959; Maia 1975; Florência 2001)
Prosodic phenomena across E. Portuguese dialects

- In this presentation we will be focusing on:
  - [j] insertion to break a hiatus
    
    \[
    a\ Aula > a\ [j]\ Aula
    \]
  - [i]/[i]-insertion at the right edge of oxitone words
    
    \[
    caf\acute{E} > caf\acute{E}[i]
    \]

- Based on speech samples collected in 3 Northern areas and 2 Southern areas
Prosodic phenomena across E. Portuguese dialects

Domains for sandhi phenomena
These sandhi phenomena are bound by the Intonational Phrase (IP)
\[ \text{domain for resyllabification} \]

- Word-final fricatives
- Word-final vowel deletion
- High V semivocalization
  \( V \ V > G \ V \)
- Syllable degemination

Compound IP
Short IP may group into a compound prosodic domain \( \rightarrow \) the higher IP is the domain for sandhi rules
Prosodic phenomena across E. Portuguese dialects

- **Domains for prosodic phenomena**

  In addition to IP, the **Prosodic Word Group** (Vigário 2010)

  Word-final e deletion:
  - Obligatory within PWG $[[...\text{e}]_{\text{PW}}...]_{\text{PWG}} \rightarrow [[...\_]_{\text{PW}}...]_{\text{PWG}}$
  - Stress clash avoidance within PWG $>$ the rule is blocked if V2 bears PWG prominence

<table>
<thead>
<tr>
<th>Example</th>
<th>Transformation</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEbe Água</td>
<td>$\rightarrow$ BEb$_\text{PW}$ Água</td>
<td></td>
</tr>
<tr>
<td>BEbe a Água</td>
<td>$\rightarrow$ BEb$_\text{PW}$ a Água</td>
<td></td>
</tr>
<tr>
<td>BEbe-a</td>
<td>$\rightarrow$ BEb$[\text{j}]_{\text{PW}}$a</td>
<td></td>
</tr>
<tr>
<td>PEde-te Água</td>
<td>$\rightarrow$ PEde-$_\text{PW}$ a Água</td>
<td></td>
</tr>
<tr>
<td>SA (Esse-Á)</td>
<td>$\rightarrow$ [Ess$[\text{j}]<em>{\text{PW}}$- [Á$\text{PW}$]$</em>{\text{PWG}}$ (Acronym=PWG)</td>
<td></td>
</tr>
<tr>
<td>um S (esse) Árabe</td>
<td>$\rightarrow$ [Ess$_\text{PW}$]$<em>{\text{PWG}}$ [Árabe]$</em>{\text{PW}}$$_{\text{PWG}}$ (Words within $\phi$)</td>
<td></td>
</tr>
</tbody>
</table>
Prosodic phenomena across E. Portuguese dialects

Sensitivity to prominence
→ Vowel deletion processes (optional) blocked under stress clash conditions

\[
\ldots o \mid PW [\dot{V} \ldots ]_{PW} \\
\ldots a \mid PW [\dot{V} \ldots ]_{PW}
\]

→ Phonological phrase (\(\phi\)) prominence, not just PW

\[
\ldots o \mid PW [\dot{V} \ldots ]_{PW} \phi \\
\ldots a \mid PW [\dot{V} \ldots ]_{PW} \phi
\]

Frota (2000)
Prosodic phenomena across E. Portuguese dialects

> e-deletion blocking if V2 is the head of PWG (not word primary stress of \(\phi\)-prominence)

- So far, only prosodic rules bound by the IP and the PWG in EP, not \(\phi\)
Prosodic phenomena across E. Portuguese dialects

→ Conditions on I formation – size, balance weight, speech rate (Frota 2000; Elordieta, Frota & Vigário 2005)
e.g. long subjects tend to be phrased as an IP; slow rate may promote a $\phi$ to IP

→ Vowel insertion phenomena:
I-final; correlated with tonal specializations → IP-internal: continuation rises > strategy of text-tune accommodation (segmental string is extended to cope with tonal realization - Frota 2000, 2002, in press; Frota & Cruz 2012)
Prosodic phenomena across E. Portuguese dialects

- Our questions:
  - What is the prosodic domain of each process?
  - Only word-level stress matters or higher levels of prominence are also relevant?
  - What is the geographical distribution of each process? (a contribution)
  - Are the processes equally active in younger and older generations? (a contribution)
Method: materials and data collection

- Data collection in loco, in a silent room
- Several taskes (at least 2h per speaker). Here, only reading task > specific corpora for the North (glide insertion to break a hiatus) and for the South (word-final epenthesis):
  - set for the North: 27 sentences (x2 per speaker)
  - set for the South: 140 sentences (x1 per speaker)
- Procedure:
  - videotaped records in .mov format;
  - audio captured by an external microphone;
  - audio extraction from the video using AoA – Audio Extractor Basic (V. 2.2.8) – output in .wav format with a sample rate of 22050Hz, mono.
Method: materials and data collection

- Speakers:
  - female
  - 3-6 per area
  - 2 age groups: 20-45 years-old and 60+ years-old
Method: materials and data collection

- Glide insertion to break a hiatus: 3 areas (Northern varieties) [Urban - U; Rural - R]
  - Viana do Castelo
  - Arcos de Valdevez (U)
  - Castro Laboreiro (R)

- Braga
- Porto
- Gião (R)

- Word-final vowel epenthesis: 2 areas in the South
  - Alentejo
  - Algarve
  - Castro Verde (U)
  - Albufeira (U)

<table>
<thead>
<tr>
<th>Area</th>
<th>No. of speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age group (rep)</td>
</tr>
<tr>
<td></td>
<td>20-45</td>
</tr>
<tr>
<td>Bra</td>
<td>2 (x2)</td>
</tr>
<tr>
<td>Erm</td>
<td>3 (x2)</td>
</tr>
<tr>
<td>Gia</td>
<td>2 (x2)</td>
</tr>
<tr>
<td>ArV</td>
<td>3 (x2)</td>
</tr>
<tr>
<td>CtL</td>
<td>3 (x2)</td>
</tr>
<tr>
<td>ALE</td>
<td>3 (x2)</td>
</tr>
<tr>
<td>ALG</td>
<td>3 (x2)</td>
</tr>
</tbody>
</table>
Method: materials and data collection

- Glide insertion to break a hiatus:
  the corpus – eliciting data with various levels of V2 prominence and various positions for V1; not aiming exhaustivity; 7 instances of stressless V2 were also introduced (resulting in 252 productions)

<table>
<thead>
<tr>
<th>V2 Promin V1 at right edge of</th>
<th>( \tilde{V} _PW )</th>
<th>( \tilde{V} _PWG )</th>
<th>( \tilde{V} _\Phi )</th>
</tr>
</thead>
<tbody>
<tr>
<td>_____ (PW-internal)</td>
<td>Simaári Cura (2</td>
<td>60)</td>
<td>________</td>
</tr>
<tr>
<td>CL</td>
<td>na á-é-gê (AEG) (3</td>
<td>90)</td>
<td>pela Ásia distante (4</td>
</tr>
<tr>
<td>PW</td>
<td>jota-á-é (JAE) (2</td>
<td>60)</td>
<td>---</td>
</tr>
<tr>
<td>PWG</td>
<td>---</td>
<td>---</td>
<td>montava asas (2</td>
</tr>
<tr>
<td>Phi</td>
<td>---</td>
<td>importava aves raras (5</td>
<td>120)</td>
</tr>
</tbody>
</table>

Total number of sentences collected per area

<table>
<thead>
<tr>
<th>Bra</th>
<th>Erm</th>
<th>Gia</th>
<th>ArV</th>
<th>CtL</th>
</tr>
</thead>
<tbody>
<tr>
<td>132</td>
<td>198</td>
<td>132</td>
<td>176</td>
<td>154</td>
</tr>
</tbody>
</table>
Method: materials and data collection

- [i]-epenthesis – the same target word in ≠ prosodic positions (35 sentences X 4 prosodic conditions):
  - internal position of the PhP ([._.]PhP) >> non-prominent
    [O café português]_{PhP} lidera. [The coffee portuguese]_{PhP} leads (the market).
  - final position of the PhP ([.._]PhP) >> prominent of PhP
    [Aquele café]_{PhP} desperta os sentidos. [That coffee]_{PhP} awake our senses.
  - final position of the IP ([.._]IP) >> prominent of IP
    [O presidente tomava esse café]_{IP}. [The president used to drink that coffee]_{IP}.
  - final position of an internal IP ([.._]IP_{min}) >> prominent of IP
    Context: [O café verde é delicioso. Sabes o que tomava o presidente?] [Tomava esse café,]_{IP_{min}} o presidente. [Used to drink that coffee,]_{IP} the president.
Results and discussion

- Glide insertion to break a hiatus

- Geographic distribution
  - [j]-insertion is attested in all areas, **to the exception of Braga (U)**
    (from now on, we will no longer consider Braga)
  - Great variation in the amount of glide insertion:
    - Erm (U) and Gia (R) (Porto) very few occurrences
    - ArV (U) and especially CtL (R) more [j] insertion.

Less or no glide closer to larger urban centers (Braga, Porto); more glide insertion in peripheral smaller village and especially in peripheral rural area
Results and discussion
- Glide insertion to break a hiatus

- The asymmetry in rate of glide insertion goes along with differences in the context where glide occurs.

Three parameters seem to be relevant:

- Type of unit V1 belongs to – PW or CL
- Prosodic domain for the rule
  \textit{Within IP, within phi, within PWG, within PW}
- Sensitivity to V2-level of prominence
  V2 head of: Phi, PWG, PW
Results and discussion
- Glide insertion to break a hiatus

- Type of unit V1 belongs to: PW or CL

- Glide insertion only applies when V1 belongs to a CL

→ Erm, Gia > glide insertion only applies when V1 belongs to a CL
Results and discussion
- Glide insertion to break a hiatus

- Prosodic domain – V1 is not part of CL

- Geographic variation in the domain of glide insertion:
  - CtL – span rule within IP
  - ArV – span rule within PWG
Results and discussion
- Glide insertion to break a hiatus

- Sensitivity to V2-level of prominence
  - All 252 instances of stressless V2 there was no [j]-insertion
  - All instances of glide insertion occurred with stress in V2 > stress matters

BUT
All levels of stress do not trigger glide insertion evenly
Results and discussion
- Glide insertion to break a hiatus

- Sensitivity to V2-level of prominence – V1 part of PW

Variation in the sensitivity to V2-level of prominence

CtL: effect of domain and level of prominence – lower level>less glide insertion; non-head>less/no glide insertion
ArV: no effect of V2-level of prominence >PW prominence
Results and discussion
- Glide insertion to break a hiatus

- Sensitivity to V2-level of prominence – V1 part of CL

→ All areas show effects of V2 prominence, but not evenly:
  CtL: 'V2_InsidePW < 'V2_PWG, 'V2_Phi
  ArV: ~NO 'V2_PW < 'V2_PWG < 'V2_Phi
  Giã, Erm: NO 'V2_PW; ~NO 'V2_PWG < 'V2_Phi
Results and discussion
- Glide insertion to break a hiatus

- Age effects

- In most areas and contexts younger subjects insert glide less (not enough data for CtL – only one 60+ speaker, 1rep; ~Giã)
- Some speakers in the lower-insertion area never inserted [j] > 3 of 5 of the younger group in the Porto area
Results and discussion

Word-final epenthesis (by segmental context)

- Only previously described contexts are attested (added segmental contexts not relevant).
- Paragoge does not occur in Alg (from now on, we will no longer consider Alg).

![Segmental contexts: known vs. new](chart.png)
Results and discussion

Word-final epenthesis (by segmental context)

- Older speakers produce + paragogic V.
- Two contexts display more insertion in both groups [r]_ and <e>_
- No insertion after [E], unlike previous reports
Results and discussion

Word-final epenthesis (by prosodic condition)

- IP prominence: relevant for the word-final epenthesis (in both age groups).
- Although not so frequently, paragoge also occurs at the IPmin boundary, which allows to conclude that the IP (and not the U) prominence triggers word-final epenthesis.
Results and discussion
Word-final epenthesis (by prosodic condition)

- Word-final epenthesis also occurs in the Standard variety (SEP), both at IPmin and IPmax heads. However, the phenomenon is not the same.

  - IP-final position: only in interrogatives and voc. chants
  - IPmin: only with continuation rises
  - strategy of text-tune accommodation: when the tune is too complex for the text, the segmental string is extended to cope with tonal realization (Frota 2000, 2002, in press; Frota & Cruz 2012)

- Ale
  - IP-final position: also in declaratives
  - IPmin: also with non-final falling contours
  - not a strategy of text-tune accommodation: nuclear contour of declaratives is not complex.
Results and discussion

Word-final epenthesis (by prosodic condition)

Neutral yes-no question produced by AG (SEP). Epenthesis as a strategy of text-tune accommodation.

Neutral declarative with a topicalized complement in situ, produced by AC (Ale, 60+). Epenthesis with an IPmin nuclear falling contour.

Neutral declarative, produced by LM (Ale, 20-45). Epenthesis is not a strategy of text-tune accommodation (nuclear pitch accent and boundary tone are both monotonal).
Main Conclusions

Glide insertion to break a hiatus

- Glide insertion was not found in Braga
  In all other areas it is found when V1 belongs to a CL
- In all these areas there is an effect of the level of prominence of V2 with CL: glide insertion appear much more/only when V2 is the head of Phi
- The effect of prominence in other contexts is only confirmed in Castro Laboreiro
Main Conclusions

Glide insertion to break a hiatus

- Castro Laboreiro and ArV are the only areas where glide insertion seems to pattern like an active, fully prosodic process, but there are three main differences:
  - In CtL the domain for the rule is higher than that of ArV (IP vs PWG)
  - In CtL it is sensitive to the level of prominence of V2 (probably) in all contexts, whereas sensitivity to V2-level of prominence in ArV is already restricted to the contexts where V1 belongs to a CL
  - The overall amount of glide insertion is much higher in CtL than in ArV, and the same considering in each context
Results suggest a pattern of rule loss/change

1 – domain of the rule becomes lower: IP > PWG > no domain/no rule, only CL > no rule

2 – loss of sensitivity to V2-level of prominence (prosodic): sensitivity > non-sensitivity

3 – sensitivity to V2-level of prominence → at higher level (CL)  
   PW < PWG, Phi → PWG < Phi → Phi

4 – (only) in the area with less glide insertion and less contexts for glide insertion (Porto) some speakers never inserted glide (3/5)

5 – younger subjects tend to less glide than older subjects in all areas and in all contexts

6 – more important cities < urban peripheral < rural peripheral
Main Conclusions
Word-final epenthesis

- Evidence for IP (both prominence and prosodic level)
  - it only occurs in Ale (not in Alg, also a Southern variety).
  - frequently produced by the older group (retraction)
  - IP-prominence (not Utt-prominence) triggers the phenomenon → also occurs at IPmin-head and non-final IP.
Acknowledgements

- To all speakers involved.
- To Bruno Neto who helped in work related to data collection
- Álvaro Iriarte and Universidade do Minho, for valuable help with the Braga recordings
- This research was developed within the following projects:

**InAPoP project** (PTDC/CLE-LIN/119787/2010), funded by Fundação para a Ciência e a Tecnologia/MCTES (PIDDAC)

**PhD project** funded by Fundação para a Ciência e a Tecnologia BD/61463/2009.
Method: materials and data collection

- Based on previous segmental studies, target words were controlled in terms of:
  - stress pattern (almost with final stress, except for the target words ending with the grapheme <e> - penult stress)
  - final segmental context ([ɐ], [ɾ], [l], and the final grapheme <e> (<ponte>, bridge) that in SEP may be produced as a schwa [ˈpɔtʰ], although vowel deletion applies frequently in this specific context – Mateus & d’Andrade 2000, Vigário 2003)
  - number of syllables (monosyllables and disyllables)
  - morphosyntactic category (verbs - <cantar> (to sing) - and nouns - <radar> (radar))

- Other segmental contexts were added: (i) low vowels ([ɔ], [a]), and the nasal vowel [ɐ]; (ii) target words with a final morphologic and non-morphologic [ʃ] (<rapaz>, boy, and <sofás>, sofas, respectively)
Results and discussion
Word-final epentheses (by prosodic condition)

- Φ-prominent target words (3 cases produced by ≠ speakers) are restructured in IP-domains, and thus considered as IP-prominent >> already shown to be possible in EP (Frota & Vigário 1996; Elordieta et al. 2003; Elordieta, Frota & Vigário 2005; Frota et al. 2007).

Target: Este princípio e]PhP pratica o bem.
Produced: Este princípio e]IP_min pratica o bem.

- Although IP_min and IP_max -heads were not simultaneously controlled in terms of final segmental contexts, there are 2 cases showing that paragogic vowels may occur at both positions at a time.