Off-line and on-line processing of ambiguity: 
the role of prosodic structure 

Sónia Frota, Marina Vigário 
& Cátia Severino 

Universidade de Lisboa 
Laboratório de Fonética (CLUL/FLUL)
1. Introduction

A central problem in the study of language is how the speech signal is segmented into word and phrase-sized units.

A way to address this issue is to examine the processing of utterances with temporary ambiguities.

Prosody is known to constrain both **lexical access** and **syntactic analysis** in ambiguous sentences. Salverda et al. (2003), Chistophe et al. (2004), Millotte et al. (2007, 2008), Li & Wang (2009)


- Intonational Phrase (IP) > general agreement on the role it plays on disambiguation (although differences are reported between expert and naïve speakers)
- The role of lower phrase boundaries and of the PW is more controversial (divergent results across studies/languages studied)
1. Introduction

• We address the role of prosodic structure in the processing of both **lexical** and **syntactic** ambiguities in European Portuguese (hereafter EP)

• Our goal is to examine the effects of the full range of prosodic constituent boundaries reported for EP in previous literature and establish which boundary(ies) may constrain disambiguation, in off-line and on-line processing.

• Four experiments:
  – Two off-line studies: Listening completion; Reading completion
  – Two on-line studies: Word detection; Eye-tracking

• **Outline:** background on EP prosodic structure > experimental materials and predictions > Off-line studies > On-line studies > Summary > Conclusion
2. EP prosodic structure

- **Prosodic Word** (Vigário 2003)
  - **Word stress** (& related processes)
  - Edge-phenomena (phonotactic constraints, many segmental processes)
  - Clipping, deletion under identity

- **Prosodic Word Group** (Vigário 2009, 2010)
  - Different level of stress (with specific effects on segmental processes)
  - Blockage of phonological processes of vowel deletion
  - Focus assignment

- **Phonological Phrase** (Frota 2000, 2009)
  - Stress strengthening in SC
  - Rhythmic constraints on vowel sandhi;
    Prominence constrains pitch accent distribution

  - Many segmental processes
  - Domain for resyllabification
  - Final lengthening, pauses
  - Left-edge strengthening
  - Domain for pitch accent distribution
  - Length affects IP-phrasing
3. Materials

Pairs of sentences with temporary ambiguity were created consisting of homophonous sequences in the target stretch. Homophonous sequences have the exact same prosodic structure, to the exception of the boundary type contrast tested. The pairs of sentences are matched for number of syllables, number of PWs and number of PhPs. Syllable onsets after the target stretch were also matched.

(Sem.pre que re.ú.ne)\textbf{PhP} (a co.mi.ssão) (sa.í.mos) (mais ce.do) (do tra.ba.lho) ‘Everytime the committe meets we get out earlier from work’

(Sem.pre que re.ú.ne)\textbf{IP} (a co.mi.ssão) (sus.pen.de) (os tra.ba.lhos) (na se.de) ‘Everytime it meets, the committee cancels work assignments in the head office’

Pairs of unambiguous sentences (controls) $>$ on-line studies

Frequency of the relevant words was checked (corpus 250,000 words)
3. Materials

2 native speakers of standard EP (expert: aware of the ambiguity/naïve: not aware) were recorded individually (in a sound proof room, with a digital recorder, 44.100Hz)

Each member of a pair of ambiguous sentences appeared in a different block, interspersed with fillers (96 test items + 24 fillers)

Control sentences were produced by the expert speaker.

Prosodic conditions:

1. No boundary/PWG: pintadel/a / pinta dela ‘painting/way she looks’
2. PW/PWG: passatempos / passa tempos ‘hobbies/let the time go’
3. PW/PhP: fita-cola / ãíta cola ‘tape/stripe sticks’
4. PWG/PhP: toalha larga (A)/largã (V) ‘large towel/towel drops’
5. PWG/IP: casa amarela/casa, amarela ‘yellow house/house, yellow’
6. PhP/IP: reúne a comissãã/a reúne, a comissãã ‘committee meets/does
3. Materials: analysis of duration patterns and intonation

- For all conditions: prominence-related cues
  Syllabic duration in target stretch (HB target-LB target, ms)

<table>
<thead>
<tr>
<th></th>
<th>C 1Str</th>
<th>1Post</th>
<th>2Str</th>
<th>2Post</th>
<th>Tot_Amb</th>
</tr>
</thead>
<tbody>
<tr>
<td>No/PWG</td>
<td>31,20</td>
<td>5,50</td>
<td>0,80</td>
<td>23,33</td>
<td>73,40</td>
</tr>
<tr>
<td>PW/PWG</td>
<td>42,50</td>
<td>-3,75</td>
<td>14,00</td>
<td>7,67</td>
<td>138,00</td>
</tr>
<tr>
<td>PW/PhP</td>
<td>76,64</td>
<td>66,50</td>
<td>-50,91</td>
<td>-13,50</td>
<td>72,00</td>
</tr>
<tr>
<td>PWG/PhP</td>
<td>78,50</td>
<td>83,17</td>
<td>-46,00</td>
<td>-25,75</td>
<td>-9,13</td>
</tr>
<tr>
<td>PWG/IP</td>
<td>78,50</td>
<td>83,17</td>
<td>-85,75</td>
<td>-23,80</td>
<td>-18,25</td>
</tr>
<tr>
<td>PhP/IP</td>
<td>66,88</td>
<td>65,25</td>
<td>-53,88</td>
<td>12,50</td>
<td>92,75</td>
</tr>
</tbody>
</table>

Pitch accents and tonal boundary (presence, distribution)

<table>
<thead>
<tr>
<th></th>
<th>1T*</th>
<th>1T%</th>
<th>2T*</th>
<th>2T%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. PWG</td>
<td>(T*)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. PW</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. PWG</td>
<td>(T*)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. PW</td>
<td></td>
<td>T*</td>
<td></td>
<td>(T%)</td>
</tr>
<tr>
<td>3. PhP</td>
<td>T*</td>
<td>(T%)</td>
<td>(T*)</td>
<td></td>
</tr>
<tr>
<td>4. PWG</td>
<td></td>
<td>T*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. PhP</td>
<td>T*</td>
<td></td>
<td>(T*)</td>
<td></td>
</tr>
<tr>
<td>5. PWG</td>
<td>(T*)</td>
<td></td>
<td>T*</td>
<td>(T%)</td>
</tr>
<tr>
<td>5. IP</td>
<td>T*</td>
<td>T%</td>
<td>(T*)</td>
<td></td>
</tr>
<tr>
<td>6. PhP</td>
<td>T*</td>
<td></td>
<td>(T*)</td>
<td>(T%)</td>
</tr>
<tr>
<td>6. IP</td>
<td>T*</td>
<td>T%</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Duration differences cue prominence level distinctions / phrase-level boundaries

global/distal cues

Tonal events cue phrasal distinctions (disambiguate conditions 3 to 6, PhP and IP)
3. Materials

- Given the properties of the different prosodic domain types just described (and the results from previous work), we predict:

1. Listeners will be able to exploit PW / PWG boundaries (unlike Li & Wang 2009 results for Chinese)

2. Listeners may detect the PhP as different from the PW(G) (but probably not as clearly as in Christophe et al 2004, Milotte et al 2007, 2008 results for French, or Name & Silva 2009 for BP)

3. Clear disambiguating role of the IP
4. Off-line study I: Listening completion task

Subjects: 24 native speakers of standard EP, university students
Materials: 96 test items (plus 16 fillers)

Members of ambiguous pairs were assigned to different blocks; 12 subjects listened to block A; another 12 to block B
Test items were cut right after the end of the ambiguous word (at a zero-crossing of the amplitude signal)

Procedure: Subjects were asked to write the listened sentence and complete it as they saw fit in a response sheet. They could listen to the item as many times as they wished (Millotte et al. 2007)

- O rolo de [fita]PW cola ‘the roll of tape’ > Verb
- O rolo de [fita]PhP cola ‘the roll of stripe sticks’ > Object NP, V modifier
- Estando [ausente]PhP o João ‘John is away’ > Subject NP
- Estando [ausente]IP o João ‘Being away, John’ > Verb
4. Off-line study I: Listening completion task

Disambiguation was achieved for conditions 3 (PW/PhP), 5 (PW/IP) and 6 (PhP/IP) [4 PWG/PhP, only Expert sp.], but NOT for word-level conditions (1 No/PWG, 2 PW/PWG)

Item effects: conditions 1-4
Length effect: IP
4. Off-line study I: Listening completion task

**Conditions 1 to 4**

**Condition 1:** No boundary/PWG

*Item effect* – item 6 (always high boundary)

*Bias* towards the lower level (late closure)

None explained by word frequency

**Conditions with IP**

*Length effect* - only long phrases triggered disambiguation (nº syllables >5 Elordieta et al. 2005)

Investigate whether these effects had their source in default processing strategies that should emerge in the preferred phrasing in silent reading >> Off-line study II
4. Off-line study II: Reading completion task

Subjects: 60 native speakers of standard EP, undergraduates at UL
Materials: 43 test items (plus 8 fillers). Total 2580 responses.

The target sequence was presented in written form with the characters evenly spaced and no spaces between words

Procedure: Subjects were asked to write the piece of sentence that was given and then complete it as they saw fit.

g o s t o d a p i n t a d e l a  p i n t a d e l a
p i n t a d e l a / p i n t a d e l a (1)

d i z s e q u e p a s s a t e m p o s
p a s s a t e m p o s / p a s s a t e m p o s (2)

o r o l o d e f i t a c o l a
f i t a - c o l a / f i t a c o l a (3)

a t o a l h a l a r g a
l a r g a (A)/l a r g a (V) (4)

c a s a a m a r e l a
c a s a a m a r e l a / c a s a a m a r e l a (5)

e s t a n d o a u s e n t e o j o a o
a u s e n t e o j o a o / a u s e n t e o j o a o (6)
4. Off-line study II: Reading completion task

Overall *preference* for the **low boundary** interpretation (mean=0.2)

Strong individual item effects in conditions 1 to 4 (high), matching the listening completion results

<table>
<thead>
<tr>
<th>Factor</th>
<th>ANOVA</th>
<th>Reading completion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prosodic condition (PC)</td>
<td>by subject</td>
<td>$F(3.73, 220,33)= 71.49$, $p &lt; .001$</td>
</tr>
<tr>
<td></td>
<td>by item</td>
<td>$F(5, 37)= 3.09, p &lt; .05$</td>
</tr>
</tbody>
</table>

*Item effects* found in Exp.1 were largely predicted from the results of the Reading Comp. Pearson correlation: *Conditions 1-4, $r = .904, p < .01$; Conditions 5 and 6, $r = .442$*

No item effects in 5-6, just a *weak length effect*: short IPs mean=0 vs. Long IPs mean=0.3

Default processing strategy (late closure), transverse to all prosodic levels
Bias toward the lower boundary found in the word-level conditions in the listening task was generalised as an overall preference in the reading task (includes conditions 5-6)

All item effects found in conditions 1-4 in the listening task were replicated in the reading task

Strong length effect found at the IP-level in the listening task was not fully matched in the reading task (only a weak effect)

The results from the offline experiments show that explicit prosody in perception successfully contradicts the default preference (late closure effect), especially if an **IP-boundary (long phrase)** is involved.
5. On-line study I: Word detection (Millotte et al. 2008)

**Subjects:** 12 native speakers of standard EP, university students

**Materials:** 27 pairs as **Targets**, a subset of Listening Completion [expert speaker’s productions, condition 5 (PW/IP) not included, only test items with long phrases in condition 6 (PhP/IP)] + 27 pairs of unambiguous sentences used as **Controls** + 30 foils (20 with the same 1st syllable as the visual target, 10 ≠) + 36 fillers. Members of ambiguous pairs, as well as controls were assigned to different blocks. 6 speakers listened to each block, repeated randomly 3 times.

**Procedure:** For each pair, each target and control were presented with appropriate and nonappropriate **visual target** (8 combinations). In each trial: presentation of the visual target (1.5 s) > blank screen (1s) > audio. Trials end with response. Interval of 1s between trials. Response: Yes (Matching visual target)/ No (Non-matching visual target) **as quickly as possible**. 15-item training with feedback on reaction time. Reaction times measured from onset of target. **SuperLabPro (4.5)**

Window of analysis: target + 1\textsuperscript{st} C + 200 ms (effect of phonological similarity in foils)
5. On-line study I: Word detection

Target - A toalhaPWG **larga** mancha a roupa branca
Control - Comprei uma camaPWG **larga** para a minha casa nova
Target - A toalhaPhP **larga** muita tinta vermelha
Control - O filho do MiguelPhP **larga** tudo por causa do futebol

Ela é larga

‘It is large’

largar

‘to drop’

**Predictions**: prosodic disambiguation – clear low/high boundary difference in response type (boundary level), by prosodic condition; possibly, boundary level difference in RT; controls (as a baseline for targets) should show faster RTs but similar responses types (ambiguity factor: control/target).
5. On-line study I: Word detection

Disambiguation in all conditions (but the last: lower boundary interpreted as high, see controls!)
RTs faster for controls.
High boundary faster RT than low boundary, for targets (prosodic cues)
Similar (not for C2), less clear.
Prosodic condition*** (PhP/IP)
Boundary level***
RTs similar results:
Ambiguity*** (Ctrls > Targets)
Amb*BL * (Targ: High>Low)

More sensitivity in word-level conditions; A IP-length effect on processing (if long → IP)?
5. On-line study II: Eye-tracking

Subjects: 19 native speakers of standard EP (2 not included: contact lenses)

Materials: Same 5 prosodic conditions as in Detection, 6 pairs (targets+controls) by condition [those that could be pictured] > 60 sentences

Procedure: For each sentence, presentation of 2 pictures side by side for 3 s > central trigger (AOI dwell time 0.7 s) > same 2 pictures with audio; left/right picture counterbalancing. Total 2040 items for analysis. 5-point calibration + validation. Tracking ratio: 81.4 – 99.8. iViewX Red System from SMI, 60Hz sampling rate.

Pilot test (SuperLab): picture-based disambiguation was assessed with 10 subjects > 6 pictures changed (3 in C6)

Overall disambiguation in all conditions > picture validation

Window of analysis: end of target+200ms (eye programming) + 400ms (speech) Holmqvist et al. (to appear)

600 window better for targets (not for controls) 14,6 > 18,5 / 37,8 ; 39,3 item effects preserved (no syntax yet)
5. On-line study II: Eye-tracking

Scan path: C4 esfregona limpa
‘clean mop’/‘mop cleans’

Measure: Proportion looking time to the Low Boundary (LB) picture (1 of 3 AOIs)
5. On-line study II: Eye-tracking

Boundary Level (BL)***
Prosodic condition*BL ** (C2, 6)
BL*Ambiguity ** (ctrs > targets)
Disambiguation for
No boundary/PWG (C1)
PW/PhP (C3)
PWG/PhP (C4)
PhP/IP (C6) [borderline]

On-line studies: summary
Word-level (C1) is disambiguated in both exps (unlike in off-line exps), but not C2:
PW/PhP (C3) is disambiguated in all listening exps
PWG/PhP (C4) is disambiguated in both exps (confirms the off-line result for expert)
PhP/IP (C6) is not clear: one item shows the IP length effect (if long > IP)

More sensitivity to word-level distinctions in on-line tasks;
but no clear support for the length effect in the processing of IPs
6. Summary and conclusion

- Listeners rely on information of different types of prosodic boundaries in both off-line and on-line processing of ambiguity.
- Prosodic cues may successfully contradict the default preference for late closure (Reading completion).
- Contrasts involving non-adjacent levels are disambiguated.
- Robust item effects across task (Reading completion) > future research

<table>
<thead>
<tr>
<th>P-structure</th>
<th>Cond.</th>
<th>Completion</th>
<th>Detection</th>
<th>ET</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. IP</td>
<td>3/4</td>
<td>✓ (long)</td>
<td>No (long)</td>
<td>(*)</td>
</tr>
<tr>
<td>3. PhP</td>
<td>2/4</td>
<td>✓ (long)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2. PWG</td>
<td>2/3</td>
<td>✓ (only Exp)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>1. PW</td>
<td>1/3</td>
<td>✓</td>
<td>No</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>1/2</td>
<td>No</td>
<td>No</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>0 / 2</td>
<td>No</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>0. No boundary</td>
<td>0 / 2</td>
<td>No</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

- Depending on the task, listeners are able to exploit word-level boundaries (contra Li & Wang 2009 results for Chinese; in line with predictions based on EP literature/actual cues in the data)
- Evidence for disambiguating role of the PhP in EP (in line with results for French and BP and actual cues in the data; but contra previous EP literature)
- Disambiguating role of the IP is confirmed (but subject to length conditions suggesting that size is a restriction used to predict phrasing)

**Different results** from off-line and on-line experiments argue for a multiple task approach to the study of prosody in disambiguation.
Acknowledgments

• Our 2 speakers and 120 subjects
• Audiences of PaPI 2009, Workshop on Prosody and Meaning, TIE4 (where parts of this research were presented)
• Research assistants Marisa Cruz and Nuno Matos
• Anne Christophe and Aoju Chen for helpful suggestions

Sponsoring:
FCT grant PTDC/LIN/70367/2006 and Project 3599/PPCDT
Project Frequence in Portuguese (frequency data obtained with FreP www.fl.ul.pt/LaboratorioFonetica/frep )
Laboratório de Fonética (CLUL/FLUL), Universidade de Lisboa
sonia.frota@mail.telepac.pt
marina.vigario@mail.telepac.pt  catiaseverino@gmail.com
References


