When prosody is (not) needed or enough: an eye-tracking study of prosodic disambiguation in European Portuguese

Cátia Severino, Marina Vigário & Sónia Frota

University of Lisbon
Lisbon Baby Lab
The processing of ambiguity in speech has been studied as a means to understand how prosodic cues may, or may not, impact speech processing, namely by constraining lexical access and syntactic processing.

Numerous studies across languages have shown that prosody provides different types of cues, from word to phrasal levels (Friederici, 2002, Millotte et al. 2008; Diley & MacAulley, 2008; Cutler et al. Cole, 2015; for review).

General support for the role of Intonational Phrase boundaries (IP).

More controversial effects of lower prosodic boundaries.

Different results for the effects of Phonological Phrase boundary (PhP) (e.g. French vs EP).

Word level boundaries less explored (as different from phrasal levels, and not focusing on particular phenomena, e.g. stress patterns, phonotactic cues and/or tones).
Phrasal Level boundaries: how prosody guides speech chunking and interpretation of utterances.

Relation between prosodic structure and syntactic category > Constituent Attachment.

“Hide 1, the rabbit 2, with a cloth”

In English:
* presence of high level boundary (2) favors high attachment.
* absence of high level boundary (= low level boundary) in (2) or high boundary in (1) favors low attachment (e.g. Watson & Gibson, 2005, Price et al., 1991).
* However, 5 year olds pursued high attachment (Tueswell et al., 1999; Snedeker & Trueswell, 2004)

Relation between prosodic structure and lexical access.

“Le chat 3, possessif” (vs chapeau)  “Le chat 4, grimpait” (vs chagrin)

In French:
* When compared to word level boundaries (3), higher level boundary (4) favors lexical access. (Christophe et al., 2004)
European Portuguese

Previous studies have shown that prosody guides the resolution of temporary ambiguity by adults.

There is a clear effect of Intonational Phrase boundaries (IP) > minimum number of syllables required (Elordieta et al. 2005; Severino, 2011)

- Offline task: only long IPs are interpreted as so (> 6 syllables)
- Online task: long PhPs are interpreted as IP boundaries.

A less robust effect of the Phonological Phrase (PhP), which is described as weakly marked in the language (Frota et al, 2010).

For disambiguation at word boundary levels (PW, PWG), online perception studies shown to be more sensitive in detecting them. (Vigário, 2003; Severino, 2011)

Contrasts between non-adjacent boundary levels favour disambiguity > bigger prosodic distance.
European Portuguese

Different tasks presented different results:

* offline tasks > higher sensitivity to boundaries in phrase-level disambiguation
* online tasks > higher sensitivity to prosodic boundaries in word-level disambiguation

Also reported in other studies:

* Relative clause attachment (Maia et al., 2007)
  * preference for high attachment in Portuguese with an offline task, showing more errors when material forced low attachment.
  * an initial advantage for low attachment in online tasks, with longer reaction times when materials forced high attachment.
* Prosody not fully controlled.

Different results with different task across and within languages (Watson et al, 2012)

Further studies of the role of prosodic boundaries in European Portuguese, in adult and children
As part of a project on the role of prosody in disambiguation in language development, 20 adults were tested with sentences with global ambiguity, following Brandt-Kobele & Hoehle, 2010, and Snedeker & Yuan, 2008, to:

- better understand the role of prosody in lexical and syntactic processing, using utterances with lexical and phrasal ambiguities
- examine the effects of word level prosodic cues and phrase level prosodic cues

**Prosodic contrast:**
- **Word Level Boundaries (low boundary vs high boundary)**
  - Stepped on umbrella
  - Her guard
  - Pizza rain

- **Phrasal Level Boundaries (PhP vs IP)**
  - Balloon that has a stick
  - Balloon with a stick
method: materials

* Sets of ambiguous sentences were created for the 2 experimental contrasts, which only differed in prosodic boundary type:

* Word Level contrast

  * Low: O Tito levou uma pisadela_{no boundary} (Tito was stepped on)
  
  * High: O Tito levou uma pizza_{pw} dela (Tito took her pizza)

* Phrasal Level contrast

  * Low: O Tito tira o balão_{php} com o pau. (Tito takes the balloon that has a stick)
  
  * Lowp: O Tito tira_{pause} o balão_{php} com o pau.
  
  * High: O Tito tira o balão, \_p com o pau. (Tito takes the balloon using a stick)

* control condition: unambiguous sentence for each set

  * O Tito anda de baloiço no recreio (Tito played on a swing in the playground)
A trained female speaker produced the utterances.

Acoustic analysis and intonational analysis were performed.
For each sentence, a picture of the described situation was created. Pilot testing for image validation was performed.

Pictures were paired and counterbalanced by side (L/R; R/L) within subjects.

2 experimental sets were built (Word level, Phrase Level), each of them with 2 blocks (version 1 and version 2). Subjects performed 1 block of each prosodic contrast and heard 1 sentence of each set of ambiguous utterances. Each subject was tested for all the boundaries.

Total blocks: 2 word level + 2 phrasal level

Event trials: 9 target + 8 controls
Task instructions were presented at the beginning of the experiment, followed by a training unambiguous set.

Testing phase started with two control trials, followed by pseudo-randomized order of trial presentation.

Each trial consisted of a gaze task, followed by a pointing task, where participants were asked to point to the picture matching the sentence only after instruction (*Aponta!*)

<table>
<thead>
<tr>
<th>Gaze</th>
<th>Pointing</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline (3s)</td>
<td>Pointing (3s)</td>
<td><em>Olha!</em> (Look!)</td>
</tr>
<tr>
<td>(image presentation)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sound Presentation (2s)</td>
<td></td>
<td><em>O Tito levou uma pizza</em></td>
</tr>
<tr>
<td>(sound offset aligned with picture offset)</td>
<td></td>
<td><em>dele.</em> (Tito took one of her pizzas.)</td>
</tr>
<tr>
<td>Test (3s)</td>
<td></td>
<td>No sound</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pointing order</td>
<td></td>
<td><em>Agora aponta!</em> (Now point!)</td>
</tr>
<tr>
<td>Pointing (3s)</td>
<td></td>
<td>No sound</td>
</tr>
</tbody>
</table>
Eye gaze was measured using SMI RED 500 Eye Tracker, with a 60hz tracking resolution, 5 points calibration, 70 cm screen distance.
Proportion of net dwell looking time at the target picture was used as the dependent variable. Only gaze task events were considered for gaze analysis.

2 AOI areas were defined on pictures (target/distractor)

Proportion of Looks graphs were plotted to inspect visual behavior over test phase time course on boundary levels.

Repeated Measures ANOVA

2 within-subject factors: phase (Baseline vs Test), boundary level (Low, Lowp, High, Control). 1 between-subject factor: experimental block (version 1 vs version 2)

When Mauchly’s test indicated that the assumption of sphericity had been violated, degrees of freedom were corrected.
Pointing to target picture was used as dependent variable. Only pointing task events were analysed.

Pointing direction was coded by researcher (online coding and offline coding check) through video.

Repeated Measures ANOVA

1 within-subject factor: boundary level (Low, (Lowp), High, Control). 1 between-subject factor: experimental block (version 1 vs version 2)

When Mauchly’s test indicated that the assumption of sphericity had been violated, degrees of freedom were corrected.
gaze results: word level

- **Significant main effects of Phase** ($F(1,18) = 50.65, p<.001$) and Boundary Level ($F(1.42, 25.48) = 5.35, p<.05$)
- **Significant interaction**, Phase*Boundary Level ($F(1.39, 25.01) = 31.64, p<.001$)
- **No significant effect of Experiment Block** ($F(1, 18) = 0.81, p=.78$)

Adults were able to identify low word boundaries, but struggled when facing high level prosodic word boundary.
Pointing results confirm gaze data: Identification of low word boundaries, but not as clear results for high level prosodic word boundaries. **Bias for whole word interpretation.**

- Significant main effect of Level \((F(2, 36) = 42.25, p<.001)\)
- No significant main effect of Experiment Block \((F(1,18) = 2.25, p=.15)\)
- No significant interaction, Level*Experiment version \((F(2, 36) = 2.25, p=.12)\)
gaze results: phrasal level

Regardless of boundary type, adults processed all stimuli as high attachment. Clear bias to interpret as IP, even when prosodic cues were expected to block it were present (Lowp)

- Significant main effects of Phase ($F(1,18) = 6.74, p<.05$) and Boundary Level ($F(3, 54) = 42.63, p<.001$)
- Significant interaction Phase*Level ($F(2.13, 38.26) = 46.73, p<.001$)
- No significant effect of Experiment Block ($F(1, 18) = 1.03, p=.324$)
pointing results: phrasal level

* Significant main effect of Level (F(3,54) = 80.42, p<.001)
* No significant main effect of Experiment version (F(1,18) <1)
* No significant interaction, Level*Experiment version (F(3, 54) = 2.01, p=.12)

Once again, pointing results follow gaze results: preference for high attachment, even when strong cues such as pause militates against it.
Subjects’ behaviour was as expected for controls, showing that they understood the task.

Bias for whole word, regardless of prosodic cues
- Confirms previous results (offline tasks with temporary ambiguity)

Bias for high attachment, even in the absence of cues or when strong cues as pause militates against it (Lowp).
- European Portuguese adults differ from English adults, but behave similarly to English children.

This result cannot be interpreted as a weight effect to balance IPs > EP does not allow an IP boundary after the verb for weight effects (Elordieta et al., 2005), unlike in Catalan (Prieto, 2005).
* Current results suggest 2 default processing strategies
  * Word processing differs from Phrasal processing

**Prosody is not enough**

* Ongoing work:
  * Look at language development data
    * Younger infants are more sensitive to prosodic cues and use them, at least at the word boundary level
acknowledgments

* All participants.
* Vasco Águas Oliveira for helping with images.
* Dr Joseph Butler & Dr Marisa Cruz.

Sponsoring:
* PhD research grant from Fundação para a Ciência e a Tecnologia (FCT/MEC), reference SFRH/BD/80991/2011
* Lisbon Baby Lab, University of Lisbon.
* PEst– Projecto Estratégico , Centro de Linguística da Universidade de Lisboa, reference UID/LIN/00214/2013