The intonation of one-word and first two-word utterances in European Portuguese

Sónia Frota & Marina Vigário

Universidade de Lisboa
Dep. Linguística / Laboratório de Fonética da FLUL (LinSe-CLUL)
http://www.fl.ul.pt/laboratoriofonetica/
1. Introduction

- Early intonational development in European Portuguese (EP) is largely unstudied.
- Prior work focused only on overall contour shape (Frota & Vigário 1993, 1995): a high proportion of falling contours between 0;6 and 2;0 (both in babbling and words); rising contours appear later and are relatively infrequent (Boysson-Bardies et al. 1984, Robb et al. 1989).
- In prior work the structural properties of the target language intonation system, or the pragmatic meaning of early utterances were not taken into account.
1. Introduction

Recent studies have reported different results on the relation between the acquisition of intonation and the development of grammar (and the lexicon):

- Adult inventory acquired before the two-word stage; tone-text alignment mastered from the beginning; pitch scaling mastered later (Catalan: Prieto & Vanrell 2007);
- Adult inventory not acquired before the two-word stage; systematic differences in peak alignment between early CS and AS; correlation with vocabulary size (Dutch: Chen & Fikkert 2007);
- Intonational development associated with the onset of word combinations; accent range at 1;06 similar to 4;00 (English: Snow 2006).
1. Introduction

- Our first goal: to describe the intonational properties of early utterances in EP in the AM framework
  - We address three questions:
    1. Is the inventory of pitch accents and boundary tones adult-like?
    2. Does the child master the alignment and scaling properties of tonal events?
    3. What does intonation tell us about other prosodic properties of early utterances (namely word stress)?
- Our second goal: to assess whether intonational development is correlated with grammatical and lexical development
2. Method

- A case study
  One monolingual child aged between 1;00 and 2;02 (L)
  Empirical database:
    - a longitudinal corpus of every other week videotape recordings of about 60 minutes each
      (inv+par; Lab. Psicolinguística, FLUL)
    - a corpus of audio recordings made on a nearly daily basis (par; Lab. Fonética, FLUL; available)

- Materials
  443 utterances (all 1 & 2 word meaningful utterances from 1;00 to 1;05; first 20 utterances from 1;06 to 1;11 and 2;02). 22 utts were unusable (poor sound quality) > 421
  Average 32,4 utts / month
  Criteria for meaningful utterances:
    1. Relation to adult word
    2. Context: appropriate use
    3. Consistency (in relevant stage)
    4. Adult confirmation (interaction)
2. Method

- **Transcriptions**

  Video DB: targets orthographically and phonetically transcribed in PHON (S. Correia & T. Costa); actual child production transcribed by 1 of the authors; utterances exported for analysis (wav format)

  Audio DB: targets and actual production orthographically and phonetically transcribed (M. Cruz & 1 of the authors)

- **Analysis**

  Perceptual analysis > utterance type and pragmatic meaning

  Prosodic analysis > prosodic transcription and main differences between child and adult patterns (choice of tonal events; alignment; scaling) Praat and SpeechStation2

  Reliability of prosodic transcription on the basis of 20 utts: 95% (nuclear contour)

  For Video DB, F0 values of H and L targets in nuclear position were measured (225 utterances)
2. Method

- **Prosodic analysis**


H+L* is the most common nuclear accent

<table>
<thead>
<tr>
<th>Labels</th>
<th>Realization</th>
<th>Context/ Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>(H) H+L* L%</td>
<td><img src="image.png" alt="Realization" /></td>
<td>Neutral declarative</td>
</tr>
<tr>
<td>(H) H*L+ L%</td>
<td><img src="image.png" alt="Realization" /></td>
<td>Topic phrase</td>
</tr>
<tr>
<td>(H) H<em>L+ L' H+L</em> L%</td>
<td><img src="image.png" alt="Realization" /></td>
<td>Focused declarative</td>
</tr>
<tr>
<td>L'+H H%</td>
<td><img src="image.png" alt="Realization" /></td>
<td>Continuation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Parenthetical</td>
</tr>
<tr>
<td>(H) H+L* L%</td>
<td><img src="image.png" alt="Realization" /></td>
<td>Wh-question</td>
</tr>
<tr>
<td>(H) H+L* LH%</td>
<td><img src="image.png" alt="Realization" /></td>
<td>Neutral yes-no question</td>
</tr>
</tbody>
</table>
2. Method

Our observations strongly suggest that the same set of contours is used in CDS (although with wider pitch range).

Requests are all low.

H*+L is used in focused declaratives and commands.

Two kinds of calling contours.

Only 17% of IP-internal stressed syll are accented.
3. Results

- Intonation and language development

<table>
<thead>
<tr>
<th></th>
<th>1;00</th>
<th>1;01</th>
<th>1;02</th>
<th>1;03</th>
<th>1;04</th>
<th>1;05</th>
<th>1;06</th>
<th>1;07</th>
<th>1;08</th>
<th>1;09</th>
<th>1;10</th>
<th>1;11</th>
<th>2;00</th>
<th>2;01</th>
<th>2;02</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLU w</td>
<td>1,06</td>
<td>1,14</td>
<td>1,10</td>
<td>1,03</td>
<td>1,18</td>
<td>1,05</td>
<td>1,20</td>
<td>1,29</td>
<td>1,26</td>
<td>1,46</td>
<td>1,25</td>
<td>1,18</td>
<td>1,06</td>
<td>1,07</td>
<td>1,67</td>
</tr>
<tr>
<td>wsize</td>
<td>1,13</td>
<td>1,37</td>
<td>1,62</td>
<td>1,40</td>
<td>1,52</td>
<td>1,54</td>
<td>1,53</td>
<td>1,72</td>
<td>1,43</td>
<td>1,63</td>
<td>1,70</td>
<td>1,88</td>
<td>nd</td>
<td>nd</td>
<td>1,80</td>
</tr>
</tbody>
</table>
3. Results

- Utterance type

<table>
<thead>
<tr>
<th>Type</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decl</td>
<td>38.95%</td>
</tr>
<tr>
<td>Foc</td>
<td>17.10%</td>
</tr>
<tr>
<td>Req</td>
<td>12.35%</td>
</tr>
<tr>
<td>Com</td>
<td>7.36%</td>
</tr>
<tr>
<td>Call</td>
<td>8.79%</td>
</tr>
<tr>
<td>Low call</td>
<td>4.28%</td>
</tr>
</tbody>
</table>

Word size >1.5
3. Results

- Utterance type

<table>
<thead>
<tr>
<th></th>
<th>Decl</th>
<th>Foc</th>
<th>Req</th>
<th>Com</th>
<th>Call</th>
<th>Low call</th>
</tr>
</thead>
<tbody>
<tr>
<td>38,95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17,10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12,35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7,36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8,79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4,28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Word size >1,5
<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:05</td>
<td>Request (multiword)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neutral yes-no question</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Command (late focus)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Command (early focus)</td>
<td></td>
</tr>
</tbody>
</table>

**Request (multiword)**

```
Qi
```

**Neutral yes-no question**

```
dá
dás(-me) ?
```

**Command (late focus)**

```
H+L* LH%
```

**Command (early focus)**

```
H*+L L%
```

**Fundamental frequency (Hz)**

<table>
<thead>
<tr>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>0.5</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
</tbody>
</table>

**Note:**

- *Qi* is the subject of the request.
- The neutral yes-no question is framed in a truncated sentence.
- The command is given in a late focus and early focus pattern.
- The fundamental frequency is represented in Hz.
Call
Low call
1:07
be´a be´a ‘Mami, Mami’

Call
Low call
1:08
´te: ´te: ‘Tito, Tito’
3. Results

- Choice of tonal events (% correct shape)

- Main deviant patterns

**AS**: Decl: H+L* L%  Req: H L* L%  Com: H*L L%  Call: (L)H* !H

**Int**: H+L* LH%

<table>
<thead>
<tr>
<th></th>
<th>1;00</th>
<th>1;01</th>
<th>1;02</th>
<th>1;03</th>
<th>1;04</th>
<th>1;05</th>
<th>1;06</th>
<th>1;07</th>
<th>1;08</th>
<th>1;09</th>
<th>1;10</th>
<th>1;11</th>
<th>2;02</th>
</tr>
</thead>
<tbody>
<tr>
<td>dec</td>
<td>lev</td>
<td>lev</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>req</td>
<td>LH</td>
<td>call</td>
<td>call</td>
<td>call</td>
<td>LH</td>
<td>call</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>com</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>call</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>int</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3. Results

- **Alignment**

An alignment based distinction:

  - H+L*: Alignment of the leading peak is not consistently adult-like until **1;09**: main pattern is <H+L* (L also tends to align later than in Adult EP)
  - H*+L: Alignment of H* is adult-like much earlier – **1;02**
`m6m6 ‘mum’

Foc
1;06

ta’ta ta’ta ‘Tata, Tata’

Decl
Foc
1;09

150  H+L *L%  H+L

4  tatá  tatá  4

100  180  260  H*+L L%

4  mamá  4

240  340  500  Fundamental frequency (Hz)

420  510  600  Fundamental frequency (Hz)
3. Results

- Accent and Stress

Stress patterns in disyllables:
Initially: level stress and stress shift
Final stress becomes stable first

**Level stress:** with different pitch accents & both final and penult
´m6´m6~ (1;05) m6´m6~ ‘mum’
´pa´pa (1;06) ´bOl6 ‘ball’

**R Stress shift:** with diff accents, including H+L* and H*+L
pa´pa (1;08) ´bOl6 ‘ball’

**L Stress shift:** mostly with H*+L
´m6m6 (1;06) m6´m6~ ‘mum’
3. Results

- Accent and Stress
  
  Our findings strongly suggest an interplay between accent and stress in acquisition
  
  Stress patterns in EP: Penult 76.4; Final 21.6; Monosyl 22.4  Frota et al. 2006
  Phrasal prominence is final
  Main phonetic correlate for word stress: duration Delgado Martins 2002
  Tone does not cue stress: pitch accent distribution is sparse (only 17% of IP-internal word stresses are pitch accented Vigário & Frota 2003)
  HL* is the most common nuclear accent
  
  Assuming Heads must bear a H tone (initial bias, de Lacy 2007) + properties of EP
  
  Accounts for (i) why stress is not straightforward, (ii) tendency to have H*T with penult stress or L stress shift, (iii) no L stress shift with HL*, (iv) initial later alignment of leading H in HL*
3. Results

- Scaling

Pitch scaling properties are not mastered from the beginning. Scaling seems to become stable before alignment (at least for declaratives - around 1;06).

Other aspects of use of accent range approach adult-like status: wider range in the focus accent (optional feature in Adult EP, frequent in the some speakers)
4. Summary

- **Choice of tonal events** (for the range of contours produced) is mostly correct as early as at 1;05; coincides with (i) use of a variety of utterance types, (ii) word size > 1.5
- **Inventory of pitch accents and boundary tones** is adult-like at 1;09 (=Catalan); coincides with lexicon size > 20
- **No early mastery of tone-text alignment** (≠Catalan, =Dutch): initially <H+L*. At 1;09 alignment similar to AEP. Initially, level stress and stress shift. After 1;09, stress patterns stable. Interplay between pitch accent and stress.
- **Scaling not mastered from the beginning**. It seems to become stable before alignment (≠Cat), for Decl. at 1;06. But there are aspects of scaling not mastered until later (!H in calls).
5. Conclusions

- Intonational development largely independent of the onset of the two-word stage (for L: 2;02). [similar to Catalan, but unlike in English or Dutch]:
  - Adult-like inventory of intonation contours
  - Appropriate use of distinct tunes for specific pragmatic meanings (e.g. Decl, Foc, Com, Req)
  - Development in the production of alignment and scaling patterns (1;09)

- Evidence that intonational development is correlated with increase in vocabulary size [as suggested for Dutch]
5. Conclusions

- Intonational development largely independent of the onset of the two-word stage (for L: 2;02). [similar to Catalan, but unlike in English or Dutch]:
  - Adult-like inventory of intonation contours
  - Appropriate use of distinct tunes for specific pragmatic meanings (e.g. Decl, Foc, Com, Req)
  - Development in the production of alignment and scaling patterns (1;09)
- Evidence that intonational development is correlated with increase in vocabulary size [as suggested for Dutch]

Thank you
References


Robb, Saxman & Grant. 1989. Voal fundamental frequency characteristics during the first two years of life. JASA 85.4, 1708-1717.


Acknowledgments

- Laboratório de Psicolinguística da FLUL
- Susana Correia, Teresa Costa, Marisa Cruz, Pilar Prieto
- Céu Viana for adapting P. Welby’s script
- Luma for her speech

Sponsoring:
FCT grant PTDC/LIN/70367/2006 and Project 3599/PPCDT
Laboratório de Fonética da FLUL / CLUL