Early Intonation 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, Prieto et al. 2008);
- 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 and prosodic phrasing?

- 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)

  **CDS:** exploratory analysis random sample of 50 utterances
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)
  For both DBs, duration of intervals or stop Cs within & between disyllabic words
2. Method

- **Prosodic analysis**


H+L* is the most common nuclear accent
2. Method

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

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

<table>
<thead>
<tr>
<th>Contour</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(H) L*+H HL%</td>
<td>Focused yes-no question Early focus (dashed line)</td>
</tr>
<tr>
<td>(H) L*+H LH%</td>
<td></td>
</tr>
<tr>
<td>H* L* L%</td>
<td>Request (multiword)</td>
</tr>
<tr>
<td>%H L* L%</td>
<td>Request (one word)</td>
</tr>
<tr>
<td>(H) H*+L L%</td>
<td>Command (late focus) Early focus</td>
</tr>
<tr>
<td>H*+L (L*) L%</td>
<td></td>
</tr>
<tr>
<td>L*+H (L*) L%</td>
<td>Early focus</td>
</tr>
<tr>
<td>(L+)H* !H%</td>
<td>Vocative chant (greeting)</td>
</tr>
<tr>
<td>(L+)H* L%</td>
<td>Low vocative chant (insisting call)</td>
</tr>
</tbody>
</table>
**CDS: exploratory analysis**

Random sample: 50 utterances; 1:00 – 1.10; several utter types; 20 utts unusable for acoustic analysis

Perceptual analysis: classification as AD-like (pitchwise) or Diff=Exp F0

Analysis of F0 contour: tonal shape, alignment, scaling (H & L targets)
CDS: exploratory analysis

Kruskal-Wallis ANOVA
Speaker: H, L, Range NS (p = .291; p = .322; p = .294)
Utterance type: H, L, Range NS (p = .7095; p = .575; p = .046)

Perceptual Scaling: H, Range > Sig (p = .0000; p = .0000)
L > NS (p = .5714)
3. Results

- Intonation and language development
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>17.10</td>
<td>12.35</td>
<td>7.36</td>
<td>8.79</td>
<td>4.28</td>
<td></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>3.895</td>
<td>17.10</td>
<td>12.35</td>
<td>7.36</td>
<td>8.79</td>
<td>4.28</td>
<td></td>
</tr>
</tbody>
</table>

- Word size >1.5
Request 1:05

Neutral yes-no question

dá  dá(-me) ?

Command (late focus)

Early focus

Command 1:05

’dá ‘give’
Call
Low call
1:07
be´a be´a ‘Mami, Mami’

(L+)H* !H%
Vocative chant (greeting)

(L+)H* L%
Low vocative chant (insisting call)

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

- Choice of tonal events (% correct shape)

- Main deviant patterns

**AS/CDS:**
- 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><strong>dec</strong></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><strong>req</strong></td>
<td>LH</td>
<td>call</td>
<td>call</td>
<td>call</td>
<td>LH</td>
<td>call</td>
<td>H*L</td>
<td>call</td>
<td>call</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>com</strong></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><strong>call</strong></td>
<td>L*H</td>
<td>^H+</td>
<td>^L</td>
<td>L*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H%</td>
</tr>
<tr>
<td><strong>int</strong></td>
<td>LH</td>
<td>L*H</td>
<td>L*H</td>
<td>L*H</td>
<td>L*H</td>
<td>H*L</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:
- nuclear accent of neutral / focused decl
  - 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

<table>
<thead>
<tr>
<th>Labels</th>
<th>Realization</th>
</tr>
</thead>
<tbody>
<tr>
<td>(H) H+L* L%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Labels</th>
<th>Realization</th>
</tr>
</thead>
<tbody>
<tr>
<td>(H) H*+L L%</td>
<td></td>
</tr>
<tr>
<td>(H) H*+L !H+L* L%</td>
<td></td>
</tr>
</tbody>
</table>

Decl
1;05
´m6´m6~
‘mum’

```
450
400
350
300
250
200

Fundamental frequency (Hz)
```

```
H <H+L* L%
mamà
```

```
0.5
```

```
1;05
```

```
H
```

```
4
```
Foc
1;06
‘m6m6 ‘mum’

Decl
Foc
1;09
ta´ta ta´ta ‘Tata. Tata’
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
- \text{m6} \text{m6}^\sim (1;05) \text{m6} \text{m6}^\sim ‘mum’
- \text{pa} \text{pa} (1;06) \text{bO16} ‘ball’

**R Stress shift**: with diff accents, including H+L* and H*+L
- \text{pa} \text{pa} (1;08) \text{bO16} ‘ball’

**L Stress shift**: with H*+L
- \text{m6} \text{m6} (1;06) \text{m6} \text{m6}^\sim ‘mum’
3. Results

- Accent and Stress
  
  Our findings strongly suggest an interplay between accent and stress in acquisition

  - Stress patterns in EP: Penult 76%; Final 22%+ Monosyl 22%  Frota et al. 2006
    
  - Phrasal prominence is final

  - Main phonetic correlate for word stress: duration Delgado Martins 2002

  - Stress is not cued by tone: 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


  - May account for (i) why stress is not straightforward, (ii) tendency to have H*T with penult stress or L stress shift, (iii) initial later alignment of leading H in HL* (no L stress shift with 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)
3. Results

- Phrasing

Initial stage: disyllabic targets, if uttered with 2 syllables, tend to be produced with one pitch accent per syllable (usually falling accent). 88.4% of such cases occur until 1;04;

44% of words uttered with 2 syllables show 2 pitch accents (n=38/87).

End of this stage coincides with the onset of disyllabic words: word size > 1.5 at 1;04

Other properties:
level stress predominates (64%)
longer duration of C2 in C1VC2V (similar to C2 between words)

E.g. Demuth & McCullough 2008

Suggests: syllable≈PW≈phrase>>PW≈phrase
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** ($\neq$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 ($\neq$Cat), for Decl. at 1;06. But there are aspects of scaling not mastered until later (!H in calls).
- **Evidence** (tonal, duration) for the construction of prosodic phrasing
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. Vocal 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