Prosodic structure in early child speech:
Evidence from intonation, tempo and coda production

Sónia Frota

Universidade de Lisboa
Prosodic structure in early child speech:
Evidence from intonation, tempo and coda production

Collaborators
Marina Vigário, Nuno Matos, Raquel Jordão
1. Introduction

- Studies on prosodic development have shown cross-linguistic differences in children’s early word production (e.g. Demuth 2006, for a review): in some languages word-minimality effects are strong (binary foot); in others monosyllabic monomoraic shapes prevail and persist (EP is in the latter group, Vigário et al. 2006).

- The prosodic properties of early one ‘word’ utterances may shed light on the prosodic status of the syllable(s)/foot they consist of: stress (Fikkert 1994), duration (Demuth & McCullough 2008), intonational properties (presence/absence of pitch accent, pitch accent/pitch contour type).
1. Introduction

- Studies on the early production of multiword combinations point to an initial phrase of successive single-word utts followed by a multiword phase proper (Behrens & Gut 2005, for a review).

- Several prosodic features have been used to characterise multiword combinations, and to discuss their prosodic status as integrated (or not) into a single phrase: stress and pitch shape (Behrens & Gut 2005, Grimm 2007), duration (namely, final-syllable lengthening: Hallé et al. 1991, Odorico & Carubbi 2003, Grimm 2007), presence and duration of pauses (Behrens & Gut 2005), intonational structure (distribution of pitch accents, pitch accent types and their combinations, boundary tones: Chen & Fikkert 2007).
1. Introduction

- Recent studies on prosodic development in European Portuguese (EP) have independently gathered production data on intonational development, word stress, pauses, and duration patterns at segmental, syllabic and phrasal levels, in early child speech (Frota & Vigário 2008a,b, Frota & Matos 2009, Frota et al. in progress)

- In the present talk, these data will be described and discussed in an integrated way.

- Our main goal: to describe the development of prosodic structure, on the basis of evidence from intonation and tempo.

- A related question also to be addressed is whether the child uses the acquired prosodic structure to constrain the development of other aspects of phonology (coda production: Jordão 2009, Jordão & Frota 2009)
2. Background

- Prosodic and intonational properties of EP (Frota, to appear)
  - Standard declarative contour: (H*) H+L* Li
  - Most common nuclear accent: H+L* (decl, wh-Q, yes-no Q; !H+L*)
  - Most common prenuclear accent across sentence types: H*
  - Relevant domain for PAD is the Intonational Phrase > sparse distribution: Only 17% of IP-internal stressed syllables are accented; accentless high plateau intermediate between the initial and nuclear accent.
  - Sentences are phrased into a single IP (but length constraints); parentheticals, topics, vocatives are phrased independently
  - Phrasal boundaries: rising or sustained (H%, !H%); Low (L%) [attain the bottom of the range in the speech event]
  - Calling and vocative contours: (L)H* !H (chanting) (L)H* L%
  - Phonological alignment contrast: H+L* / H*+L (focus, commands)
Long subject

Topic

Two root sentences
2. Background

- Prosodic and intonational properties of EP
  - As in many other languages, phrase-final lengthening. However, only at the IP-level (Frota 2000)
  - Duration is a cue to word stress (Andrade & Viana 1989, Delgado Martins 2002)
  - Duration also cues the difference between strong and weak forms of clitics at IP-initial position (Frota 2000)
  - Pauses are inserted at IP-boundaries (Frota 2000, Cruz 2009, Serra 2009)
2. Background

- Early intonation in EP (Frota & Vigário 2008a,b)
  - Diversity in nuclear contours at 1;05
  - Inventory of nuclear accents and boundary tones is adult-like at 1;09 (similar to Catalan - Prieto & Vanrell 2007); coincides with lexicon size > 20)
  - Development in the production of alignment and scaling patterns (1;09)
  - Intonational development largely independent of the onset of the two-word stage (2;02)
  - Intonation, stress and segmental duration data suggested a path of prosodic development in three phases: syllable≈PW≈phrase>> syllable≠PW≈phrase>> syllable≠PW≠phrase

This is the hypothesis to be tested with more data, of various kinds, focusing on the analysis of prosodic structure.
3. Method

- **A case study**
  One monolingual EP child aged between 1;00 and 2;04 (Luma)

- **Empirical database** (recording density and combination of different kinds of records):
  - a longitudinal corpus of every other week videotape recordings of about 60 minutes each
  - a corpus of audio recordings made on a nearly daily basis
  - a detailed parental diary

[the latter two are available: LumaLiDa Resources, Lab. Fonética]

- **Transcriptions**

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

**Audio DB:** targets and actual production orthographically and phonetically transcribed (M. Cruz, N. Matos & the author)
3. Method

- **Materials**
  
  **Intonational analysis:**
  
  506 utterances (all 1 & 2 word meaningful utterances from 1;00 to 1;05; first 20 utterances from 1;06 to 2;01; first 20 multiword utts from 2;02 to 2;04). 22 utts were unusable (poor sound quality) > 484

  Average 30 utts / month

  Criteria for meaningful utterances:
  1. Relation to adult word
  2. Context: appropriate use
  3. Consistency (in relevant stage)
  4. Adult confirmation (interaction)

- **Analysis**
  
  **Any** one-word and multiword combinations within a single speech event

  Perceptual analysis > utterance type and pragmatic meaning [Context]

  Prosodic analysis > prosodic transcription: **pitch accent type & distribution + boundary tones. Reset.**

  Downstep/upstep relations. **Presence of pauses (>300ms next to stops).** Praat and SpeechStation2

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

  For both DBs, **duration of intervals or stop Cs within & between disyllabic words**
3. Method

- **Materials**

  Analysis of syllable durations:
  
  334 utterances (distributed into 6 periods – see Table) from the Audio Database

  The same criteria for meaningful utterances applied:

  - **Words** – identified relative to target and context
  - Utterances - “vocalizations separated from all others by audible breaths or in accord with adult judges’ intuitions about utterance boundaries” (Oller & Linch (1992): IP

- **Analysis**

  Acoustic measurements followed standard criteria (Turk et al. 2006)
  
  - SpeechStation 2.0: Spectrograms + waveforms
  
  Status of each syllable > stress and position in PW and IP
  
  (initial, medial, final, monosyllabic)

  - **Size of the PW** > nº syllables
  - **Word duration**
  - **Size of the IP** > nº syllables; nº PWs

  Frota & Matos 2009; Matos, in progress

<table>
<thead>
<tr>
<th>Age</th>
<th>1;01</th>
<th>1;04</th>
<th>1;06</th>
<th>1;08-09</th>
<th>2;02</th>
<th>2;03-04</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>124</td>
<td>101</td>
<td>46</td>
<td>179</td>
<td>94</td>
<td>212</td>
<td>756</td>
</tr>
<tr>
<td>PW</td>
<td>84</td>
<td>79</td>
<td>39</td>
<td>122</td>
<td>75</td>
<td>146</td>
<td>545</td>
</tr>
<tr>
<td>Utt</td>
<td>35</td>
<td>71</td>
<td>29</td>
<td>74</td>
<td>57</td>
<td>68</td>
<td>334</td>
</tr>
</tbody>
</table>
3. Method

Luma’s language development

|       | 1:00 | 1:01 | 1:02 | 1:03 | 1:04 | 1:05 | 1:06 | 1:07 | 1:08 | 1:09 | 1:10 | 1:11 | 2:00 | 2:01 | 2:02 | 2:03 | 2:04 |
|-------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| MLUw  | 1.06 | 1.14 | 1.10 | 1.03 | 1.18 | 1.05 | 1.20 | 1.29 | 1.26 | 1.46 | 1.25 | 1.18 | 1.06 | 1.07 | 1.67 | 1.58 | 2.06 |
| wsize | 1.13 | 1.37 | 1.62 | 1.40 | 1.52 | 1.54 | 1.53 | 1.72 | 1.43 | 1.63 | 1.70 | 1.88 | nd   | nd   | 1.80 | 1.65 | 1.55 |
4. Results: intonational analysis

- Production of disyllabic targets (initial stage)

  If not truncated, tend to be produced with one pitch accent per syllable (usually falling accent): until 1;04

  Level stress predominates (64%)

  The intervocalic stop consonant is realized with a duration longer than its counterpart in productions showing just one pitch accent and similar to the duration of the stop interval between two target words
Further, the fact the syllable may be uttered with different nuclear contours correlated with different pragmatic meanings (decl, request, calling – Frota & Vigário 2008b) shows that it functions as a prosodic phrase.
ta’ta ‘Tata’ (the way the child calls herself)

These findings strongly suggest that the syllable is being treated as a PW.

The end of this stage (syllable≈PW≈phrase) coincides with the onset of disyllabic words: word size > 1.5 at 1;04
4. Results: intonational analysis

- Production of disyllabic targets (after the initial stage)
  The analysis of the next 4 months shows that one pitch accent per syllable is only a residual strategy.
  Truncation is also reduced.
  One pitch accent per word: 80%.
  Patterns of stress shift predominate over level stress.
  These facts indicate that syllables coincide no longer with PWs:
  
  `k6`t6 (1;01) `t`a`t`a child’s name
  `p`a`p`a (1;03) `bO`l6 ‘ball’
  `m6m6 (1;06) m`6`m6~ ‘mum’
  `p`a`p`a (1;08) `bO`l6 ‘ball’

  syllable ≠ PW≈phrase
4. Results: intonational analysis

- Production of multiword combinations within a single speech event

When and how PWs become integrated within the same phrase?

Analysis: pitch accentuation, presence of pause and reset between any 2 words, pitch accent combinations

Stage 1 until 1;04; (2) 1;05 until 1;09 (previous results on intonational development; lexicon size >20);
(3a) 1;10 until 2;01 and (3b) 2;02 onwards (MLU>1.5).

1: n=11; 2:n=14; 3a: n=20; 3b: n=55

- Pitch accentuation and reset indicate that prosodic integration occurs after 1;09
- Presence of pauses suggests that temporal integration is more gradual
- The occurrence of combinations of diff pitch accents shows a difference between 1 (repetition of pitch pattern), 2 (different patterns) and 3 (equal dist.)
Until 1;09: All cases show pauses and pitch reset. 1st example shows repetition of pitch pattern, typical of stage 1 (=1PA per syllable cases). Other examples show different patterns (predominant pattern 1;05-1;09). 1;05-1;09: sequences of calling contours, sequences of vocative(calling)+declarative/imperative (1 case V+N).
Top: After 1;09, illustrative examples of 1st cases of prosodic integration: prenuclear H* appears; contour shows a high plateau or downstep. Integration is not syntactically motivated: vocative+V; topic+Adv.
Bottom: in 3-word combinations, N(Subj) is phrased alone (59%). Cases of higher level integration (sentence level) are rare (18%). Ex. of accentless word ([a] VERB).
4. Results: intonational analysis

Unaccented words:
all prenuclear, monosyllables (88%)
Prenuclear H* is dominant
(from 2;02 onwards: 53%)
Cases of prosodic integration not motivated by syntax continue to appear after 2;02 (voc+X; interjection+noun); cases of syntactic phrases not integrated also occur (N+Adj).
When and how PWs become integrated within the same phrase? After 1;09, independently of syntax.

- Interim Summary
Findings support 3 developmental phases:
Stage 1 (until 1;04) [word size]
syllable≈PW≈phrase
Stage 2 (until 1;09) [lexicon size]
syllable ≠ PW≈phrase
Stage 3 (1;10 onwards)
syllable ≠ PW ≠ phrase
4. Results: temporal analysis

• Correlations between duration and nº of units within a domain (Pearson), and their development:

Results suggest two moments of reorganization of temporal patterns:

- 1;04 – PW duration and nº of syllables: PW duration until 1;04, and unlike in later production, is not correlated with number of syllables per PW
- 1;08-09 – syllable duration and nº of units at higher domains

**Dur syl/nºsyl IP:** a temporal reorganization occurs at the phrasal level

Development of correlation patterns between syllable/word duration and nº of units at higher levels > ‘**u-shape**’ – discontinuity pattern explained by linguistic acquisition vs. Continuity based on biological factores (Robb & Saxman 1990, Snow 1994, 2006) > **emergence of prosodic structure**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1;01</td>
<td>Wp10-20a</td>
<td>-.624**</td>
<td>-.307**</td>
<td>-.192*</td>
<td>-.222*</td>
</tr>
<tr>
<td>1;04</td>
<td>wp10-20b</td>
<td>-.397**</td>
<td>-.291**</td>
<td>.170</td>
<td>.067</td>
</tr>
<tr>
<td>1;06</td>
<td>wp10-20c</td>
<td>.177</td>
<td>-.378**</td>
<td>.657**</td>
<td>-.355*</td>
</tr>
<tr>
<td>1;08-09</td>
<td>wp&gt;20</td>
<td>-.186*</td>
<td>-.080</td>
<td>.534**</td>
<td>.083</td>
</tr>
<tr>
<td>2;02</td>
<td>wp&gt;50a</td>
<td>-.336**</td>
<td>-.465**</td>
<td>.312**</td>
<td>-.266**</td>
</tr>
<tr>
<td>2;03-04</td>
<td>wp&gt;50b</td>
<td>-.251**</td>
<td>-.294**</td>
<td>.302**</td>
<td>-.196**</td>
</tr>
</tbody>
</table>
4. Results: temporal analysis

- Syllable duration by position
  - Up to 1;09, **similar pattern** of durations in PW and IP
  - From 1;09 onwards, **clear final lengthening at IP-level** and reduction of duration of syllables at I and M positions
  - Supports **reorganization** at 1;08-09 (correlations)
  - **Separation PW / IP:** final lengthening at IP-level in adult speech vs. PW (Frota 2000)

In short, duration findings point to the same 3 stages
Additional evidence: coda emergence

Intonation and duration facts show that PWs and prosodic phrases are treated differently after 1;09 (phase 3).

Coda development data gives further evidence: codas start to be treated at IP prominent and IP final positions, and not at PW prominent or final positions.

**Repair strategies**

Luma: after 2;00 (V epenthesis)

The PW / prosodic phrase distinction is at place and constrains the development of other aspects of phonology.

Jordão 2009, Jordão & Frota 2009
5. Summary

- We discussed evidence for 3 developmental phases in the acquisition of prosodic structure in early child speech:

  Initial phase
  
  IP
  |                 
  PW
  |                 
  σ

  Evidence: truncation & 1 pitch accent per syllable (80%), level stress, C2 duration, word duration not correlated with number of syllables per word, presence of different nuclear contours with different pragmatic meanings. Until 1;04 [word size]
5. Summary

- We discussed evidence for 3 developmental phases in the acquisition of prosodic structure in early child speech:

  Initial phase > Second phase

  \[
  \begin{array}{c|c|c}
  \text{IP} & \text{IP} \\
  \hline
  \text{PW} & \text{PW} \\
  \hline
  \sigma & \sigma & \sigma
  \end{array}
  \]

  **Evidence:** 1 pitch accent per word (80%), stress shift, word duration is correlated with nº of syllables per word; nº pitch accents= nº words, pitch reset and pauses highly frequent, position of the \( \sigma \) in the IP is irrelevant for \( \sigma \) duration. _Until 1;09 [lexicon size, tonal inventory]_
5. Summary

We discussed evidence for 3 developmental phases in the acquisition of prosodic structure in early child speech:

- Initial phase
- Second phase
- Third phase

Evidence: accentless words emerge, pitch reset (and pauses) much less frequent, high plateau contour, occurrence of prenuclear H*, emergence of phrase-final lengthening

From 1;10 onwards [before MLU >1.5]
5. Discussion

- Cross-linguistic implications (universal/language-specific?):
  - May languages vary in the lower unit they chose for phase 1? E.g. Syllable vs. Foot (EP, French / English, Dutch, Japanese) Or is there an additional foot≈PW≈phrase phase for some languages?
  - The PW≈phrase stage seems to find support from observations in many languages relative to SWU or SSWU
  - However, the beginnings of prosodic integration of words, the kinds of evidence for that integration, and its relation to syntax show contradictory evidence in the literature. Our findings show that phrasal prosody emerges independently of syntax.

- Views of prosodic development (bottom-up, top-down?)
  - Key prosodic domains (lower and higher) are present and align in phase 1. Development proceeds by unfolding of the different levels. And unfolding proceeds bottom-up.
Acknowledgments

- Laboratório de Psicolinguística da FLUL
- Susana Correia, Teresa Costa, Marisa Cruz, Cátia Severino
- Luma for her speech

Sponsoring:
FCT grant PTDC/CLE-LIN/108722/2008
Laboratório de Fonética & Baby Lab, CLUL
References


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

