Prosodic and frequency effects on the development of syllable structure in European Portuguese

S. Frota*, M. J. Freitas*, M. Vigário+ & F. Martins^

*Universidade de Lisboa – FL and ONSET-CEL,
+ Universidade do Minho – ILCH,
^ Universidade de Lisboa – FL and ILTEC
1. Introduction

• Background
  – Acquisition of syllable structure in EP (overview)
  – The impact of frequency on language development
  – Frequency studies for EP
• Goals: establish the effects of syllable type frequency and the role of prosodic prominence in syllable development in EP
• Data: input (ADS & CDS), and child data
• Input frequencies and syllable types in child speech
• Prosodic prominence effects: word edges, stress
• Summary of main findings and discussion
2. Background: Acquisition of syllable structure

- **Onsets** (Freitas 1997): C and Ø > CC
  - non-branching onsets (C or Ø) are represented in the system from the beginning of production
  - empty onsets are used as a default structure to deal with problematic target onsets
  - G may replace a segmentally problematic C target
  - branching onsets are the last structure to be acquired
  
  CV, V, ØV, GV > CC

- **Rhymes** V > (surface VG>)
  - non-branching nuclei are, as expected, represented in the system from the beginning of production
  - surface VG structures are present since early stages of production, but the mastery of branching nuclei occurs in late stages of development, by the time syllable-final liquids are acquired; V and VG seem to be interpreted similarly by children (Correia, 2004)

  V, VG > VC
2. Background: Acquisition of syllable structure

- Children early productions (examples)

<table>
<thead>
<tr>
<th></th>
<th>Production</th>
<th>Transcription</th>
<th>Duration</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>pato /'pa.tu/ → ['tɐ] (João: 0;11.6)</td>
<td>‘duck’</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>dá /'da/ → ['da] (Inês: 0;11.14)</td>
<td>‘give’</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>quer /'kɐɾ/ → ['kɐ] (Marta: 1;2.0)</td>
<td>‘wants’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>água /'a.gwɐ/ → ['a.βɐ] (João: 0;11.6)</td>
<td>‘water’</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>é /'e/ → ['e] (Inês: 1;0.25)</td>
<td>‘is’</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>água /'a.gwɐ/ → ['a.wɐ] (Marta: 1;2.0)</td>
<td>‘water’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>vês /'vɐʃ/ → ['ɐʃ] (Marta: 1;3.8)</td>
<td>‘see’</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>flor /'flɔɾ/ → ['ɔli] (Inês: 1;9.19)</td>
<td>‘flower’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td>quer /'kɐɾ/ → ['kɐ.ɾi]/['kɐ.ɾi] (Inês: 1;10.29)</td>
<td>‘wants’</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>papel /pɐ.'pɐ̃.l/ → [pɐ.'pɐ̃.l] (Marta: 2;2.17)</td>
<td>‘paper’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e.</td>
<td>pato /'pa.tu/ → ['ɐ.'tɐ]/['tɐ] (João: 0;11.06)</td>
<td>‘duck’</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>chupeta /ju.'pe.tɐ/ → ['ɐ.'pi]/['pi] (Inês: 1;01.30)</td>
<td>‘pacifier’</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>dá /'da/ → ['ɐ.'da]/['da] (Inês: 1;0.25)</td>
<td>‘give (me)’</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cão /'kẽw/ → ['ɐ.'kẽw]/['kẽu] (Marta: 1;02.0)</td>
<td>‘dog’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2. Background: the impact of frequency

- There is a growing interest in determining the importance of frequency information in linguistic behaviour (Bybee, 2000 & 2001; Bybee & Hooper, 2001; Jurafsky, Bell & Girand, 2002; Moates, Bond & Stockmal, 2002; Pierrehumbert, 2002)

- The same is true for the acquisition and development of linguistic systems (Fikkert & Freitas, 1998; Lleó & Demuth, 1999; Beckman e Edwards, 2000; Roark & Demuth, 2000; Demuth & Johnson, 2003; Prieto, 2004; Vigário, Freitas & Frota 2005)

- Frequency in EP (adult speech): syllable types (Andrade & Viana, 1994; Vigário & Falé, 1994; Viana et al., 1996); phonetic segments and stress patterns (Viana et al., 1996); words undergoing phonetic reduction (Vigário, 2003); *minimal* words and patterns of cliticization (Vigário, Martins & Frota, 2005)

- Frequency in EP (child speech): development of syllable structure (Fikkert & Freitas; Vigário, Freitas & Frota, 2003); development of word shapes (Vigário, Freitas & Frota, 2005)
2. Background: the frequency tool \textit{FreP}

- \textit{FreP} (Vigário, Martins & Frota, 2005)
  - Automatically extracts syllable types from written texts;
  - Semi-phonological: includes lexical phonological processes (e.g. glide insertion to break a hiatus); ignores all optional processes (deletion of unstressed vowels, optional gliding);
  - Includes glides in rising diphthongs that are obligatory (post-tonic in proparoxitons) and V positions between consonants that violate syllable construction principles (Mateus & Andrade 2000);
  - Glides between vowels are treated as ambisyllabic

- **Goals**
  - Establish the EP-specific frequency distributions of syllable types in the input
  - Measure the effect of input frequencies on syllable type development
  - Assess the role of prosodic prominence:
    - position in the word: \textit{word-edges/monosylw > word internal}
    - stressed > unstressed
  - Discuss the impact of frequency on development
3. Data (number of syllables by syllable types)

- **Spontaneous data from 3 monolingual Portuguese children (CS):**
  - João aged 0;10.2 – 1;8.13 (n=1003)
  - Inês aged 0;11.14 - 1;10.29 (n=3619)
  - Marta aged 1;2.0 - 2;2.17 (n=6090)  
  (Freitas 1997)

- **Child-directed speech data (CDS):**
  - Inês (n=24867)
  - Marta (n= 10985)

- **Spontaneous Adult data (ADS):**
  - the *Português Falado* corpus (CLUL – CDRom) from 90s (n=41826)
  (Vigário, Martins & Frota 2005)

Data obtained with *FreP* (Reliability was above 99.5%)
4. Input frequencies

CV, V, CVC = 73%

CV>V>(C)V>(C)V>(C)VG(N)>CCV
4. Input frequencies

CV, V, CVC = 78%

CV>V>(C)VC>(C)VN/(C)VG(N)>CCV

CDS more V (+5%)
less types (25 vs. 29)
less freq. lower %
4. Development of syllable types

João - Syllable types

CV>V>CVC>CVG/CVN>CCV
CV / V > (C)VN > (C)VG, GV
4. Development of syllable types

Inês - Syllable types

CV > V > CVC > CVG/CVN > CCV
CV / V > (C)VN > (C)VG > (C)VC

![Graph showing the development of syllable types with specific labels for each type and age intervals.](image)
4. Development of syllable types

Marta - Syllable Types

CV>V>CVC>CVG/CVN>CCV
CV / V , (C)VN, (C)VG > (C)VC
5. Prosodic Prominence and frequency

- **Two mismatches**
  - Development of syllable types and input frequencies: CV and V
    (C)VG/(C)VN before (C)VC
  - Hypothesis: prosodic prominence (stress and/or word-edges) may have a role to enhance V, and CVG/N relative to CVC in the input
    e.g. Echols 1987; Peters 1977, 1983

- **Prosodic Prominence (EP)**
  - Word-edges (Vigário 2003):
    Prosodic word initial position
    - often gets emphatic stress
    - vowels are exempt from regular reduction in stressless positions
      e.g. [e / ɐ]rguEr vs. ro[i]dOr
      [o / ɐ]piniÃo vs. mi[u]lInho
    Prosodic word final position
    - tends to show more variable/complex structures (more V reduction/deletion)
  - Stress: stressed syllables are acoustically longer; no vowel reduction
5. Input: word-edges vs. internal position

Syllable diversity & complexity:
Final > Monow > Initial > Internal

ADS - Syllables by position in w
5. Input: word-edges vs. internal position

CDS - Syllables by position in w

Syllable diversity & complexity:
Final > Monow > Initial > Internal

2 speakers = patterns
CDS more V
5. Word-edges vs. internal position

**Input**
- V syllables in ADS, CDS
  - Initial+monopw: 54%, 56%
  - Initial+monow: 90%, 95%
  - V clitics are proclitics and occupy initial position (don’t reduce if followed by C and never delete – [u] carro; [u / w] aluno)
  - Initial V syllables are not reduced even if stressless
  - V syllables in monopw are obviously stressed

*Most V syllables occur in a prominent word position*

**Input**
- CVG/N and CVC in ADS, CDS
  - Final+monopw
    - CVGN: 88%, 100%
    - CVC: 62%, 74%
  - Final syllables and monosyllabic words display more diversity and complexity of types
  - CVGN and CVC in monopw are obviously stressed

*Most CVGN and CVC syllables occur in a prominent word position, but CVGN does so more often (mainly in CDS)*
5. Word-edges vs. internal position

• CS: CV/V > (C)VN > (C)VG > (C)VC
  - So, prosodic prominence may have a role to enhance V, and CVG/N relative to CVC in the development of syllables types
  - Initial w-position displays more V syllables than all other positions (as in the input but with higher frequency)
  - Syllable diversity and complexity in CS closely mirrors the input: Final, Monow>Initial>Internal

• CS
  - More complex types appear first in final w-position and monosyllabic words and are more frequent in these positions

<table>
<thead>
<tr>
<th>Types</th>
<th>INITIAL</th>
<th>FINAL</th>
<th>MONOW</th>
<th>INTER</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV</td>
<td>BEGIN</td>
<td>BEGIN</td>
<td>BEGIN</td>
<td>BEGIN</td>
</tr>
<tr>
<td>V</td>
<td>BEGIN</td>
<td>BEGIN</td>
<td>BEGIN</td>
<td>BEGIN</td>
</tr>
<tr>
<td>(C)VG</td>
<td>J, I, M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>(C)VN</td>
<td>J, I, M</td>
<td>J, I</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>(C)VGN</td>
<td>J, I, M</td>
<td>I, M</td>
<td>J, I, M</td>
<td>------</td>
</tr>
<tr>
<td>(C)VC</td>
<td>M</td>
<td>I, M</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>(C)VGC</td>
<td>------</td>
<td>I</td>
<td>I, M</td>
<td>------</td>
</tr>
<tr>
<td>CC...</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5. Word-edges – initial position

Inês - Initial Position in w
5. Word-edges – final position

Inês - Final Position in w
5. Word stress

**ADS**
- We saw that word position **may have** a role to enhance V, and CVG/N relative to CVC in the development of syllables types. What about stress?
- CV, V, CVC more frequent in **unstressed** position
- CVG, CVN more frequent in **stressed** position
  - CVN 82%; CVG(N) 66%

**Stress strengthens the word position effect**

**CDS**
- CV, V more frequent in **unstressed** position
- CVG, CVN and CVC more frequent in **stressed** position
  - CVG(N) 98%
  - CVN 78%
  - CVC 60%
- Still CVG and CVN occur **more frequently** in stressed position!

**Overall, prosodic prominence may play a role to enhance CVN/G relative to CVC**
5. Word stress: CV, V

ADS - The effect of stress
### 5. Word stress: CVN, CVG

**ADS - The effect of stress**

<table>
<thead>
<tr>
<th>Position</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVN</td>
<td>6</td>
</tr>
<tr>
<td>CVG</td>
<td>5</td>
</tr>
<tr>
<td>CVGN</td>
<td>4</td>
</tr>
<tr>
<td>VN</td>
<td>3</td>
</tr>
<tr>
<td>CVC</td>
<td>2</td>
</tr>
<tr>
<td>VN</td>
<td>1</td>
</tr>
<tr>
<td>VC</td>
<td>0</td>
</tr>
</tbody>
</table>

The graph illustrates the frequency of stress positions for CVN, CVG, CVGN, VN, CVC, and VC. The x-axis represents the position (stressed or unstressed), and the y-axis represents the frequency (%) of each stress pattern.
6. Summary

- EP frequency distributions of syllable types in the input (ADS and CDS) predict the following order of emergence of syllable types: CV > V > CVC > CVG/CVN > CCV
- CS: CV / V > (C)VN, (C)VG > (C)VC
- Two mismatches: CV and V
  (C)VG/(C)VN before CVC
  - Role of prosodic prominence offers an explanation to the mismatches: word-edges vs. internal position stressed vs. unstressed position
- ADS & CDS: Most V syllables occur in word-initial position
- ADS & CDS: Most CVGN syllables occur in final position and CVGN outranks CVC in this position
- ADS & CDS: CVG(N), CVN more frequent in stressed position

✓ Structural information (prosodic prominence) and frequency together predict the correct order of emergence of syllable types

CV > V > CVC > CVG/CVN > CCV
CV / V > CVG/CVN > CVC > CCV
Prosodic and frequency effects on the development of syllable structure in European Portuguese

S. Frota*, M. J. Freitas*, M. Vigário† & F. Martins^

*Universidade de Lisboa – FL and ONSET-CEL,
† Universidade do Minho – ILCH,
^ Universidade de Lisboa – FL and ILTEC

sonia.frota@mail.telepac.pt; joaofreitas@fl.ul.pt;
marina.vigario@mail.telepac.pt; fmartins@fl.ul.pt
• Our data: ‘semi-phonological’ syllables
• Vowel deletion [i, u]
  - Promotes surface C clusters
  - Demotes CV
• Against the syllable type patterns shown in CS
• But:
  - V deletion only in internal and final position (not in initial position)
  - V deletion only in unstressed position

• And:
  - Hypothesis: properties in the signal that cue the type of syllabic grammar, namely rhythmic properties that place EP in the group of syllable timed languages
  Vigário, Frota & Freitas 2003