5th Workshop on Phonological Development, Postdam, March 31, 2006

Early prosodic words in European Portuguese

[*] Sónia Frota

[*]Maria João Freitas

[+]Marina Vigário

[*]Universidade de Lisboa , [+]Universidade do Minho

1.Introduction

- We examine the acquisition of PWs in European Portuguese (EP), through analysis of grammatical and statistical properties of the target language & child speech

1. The relevance of the **grammar of the target system** (i.e. the phonological system) to the development of PWs
2. The impact of word shape **frequency** in the input on the emergence of early polysyllabic words and subminimal words in child speech
3. The effect of **language-specific traits** on the timing and course of PW development

2. Input Grammar: The PW in EP

- Cross-linguistic differences in PW structure:
  - ✓ Ls showing a constellation of phenomena cuing the PW: Dutch, German, English (e.g. Booij 1995, 1999; Wiese 1996, Hall 1999; Raffelsiefen 1999)
  - ✓ Ls with weaker evidence for the PW: Italian, Spanish, Brazilian Portuguese, French (e.g. Kleinhenz 1997, Peperkamp 1997, Bisol 2000, Hannahs 1995a, 1995b)
- **EP phonology** offers a **rich array of evidence** for the PW (though resyllabification is Romance-like: Vigário 2003)
- Given the solid cues to the PW, EP children may be aware of them > child speech
- If so, it's reasonable to assume that **grammar** may play a role > facilitate **early segmentation & production of the PW shapes** found in the language (e.g. Cutler 1996, Peters & Strömqvist 1996, Demuth 1996)

2. Input Grammar: (other) relevant information

- Vowel system
  - Stressed system: i, e, ɛ, a, ɔ, o, u
  - Unstressed system: i, ɨ, ɐ, u
  - No quantity contrasts
- Syllable structure: CV-dominant + V and CVC
- Word stress
  - 3 syllable window, right edge; penult stress is dominant
  - No sensitivity to quantity: [tuˈnɛɫ]/[ˈtunɛɫ] 'large cask/tunnel'
    [tuˈfɐ̃w̃]/ [ˈɔɾfɐ̃w̃] 'tornado/orphan'
- PWs (lexicon & most frequent words): 1 to >3 syllables, monosyllabic with open syllables included
  - [ˈpɛ] 'foot'; [ˈmaɾ] 'sea'; [tɾɐ.bɐ.ʎɐ.ˈdo.ɾɐʃ] 'workers (fem)

2.1. Input Grammar: ***Evidence for PW edges***

- Left-edge
  - Phonotactic constraints *[ʎ/ɲ/ɾ/ɨ…]$_{\omega}$

[**ʀ**]ato vs. ca[**ɾ**]o 'mouse'/'expensive'

*[**ɨ**]rguer vs. p[**ɨ**]rder 'raise'/loose'

  - Non-central vowels do not undergo full vowel reduction

[**e**]rguEr vs. ro[**ɨ**]dOr 'raise'/'rodent'

[**o**]piniÃo vs. mi[**u**]lInho

'opinion'/ 'soft part of bread-DIM'

  - Emphatic stress (optional) high F0 level

- Right-edge
  - Stressless vowels in σs closed by sonorants are lowered

lÍd[**ɛ**]r, Âmb[**a**]r vs. pOd~~e~~, acAb[**ɐ**]

'leader' 'amber' 'can' 'finish'

  - Stressless Vs ending ωs within morphological compounds are lowered:

mOn[**ɔ**]-acentuAl vs. mon[**u**]grafIa

'mono-accentual' 'monograph'

  - Non-high palatal Vs are deleted

pass~~e~~ vs. pass[**e**]mos, pass[**j**]ar

'pass' 'pass-SUBJ-2pp' 'take a walk'

2.1. Input Grammar: *Evidence for PW*

- PW-bound phenomena
  - –Two processes of PW-deletion
    - clipping

(tele)$_{\omega}$(móvel)$_{\omega}$ > móvel ‘mobile’ vs.
(telefonia)$_{\omega}$ > *fonia ‘radio’

- deletion under identity

(mono)$_{\omega}$(gamia)$_{\omega}$ ou (poli)$_{\omega}$(gamia)$_{\omega}$ >
mono ou poligamia ‘mono/polygamy’
vs. (biografia)$_{\omega}$ e (discografia)$_{\omega}$ >
*bio e discografia
‘biography and discography’

(Vigário 2003)

- Prominence-related phenomena
  - – Word stress
    - perceptually salient
    - many segmental rules referring to the presence/absence of stress

e.g. vowel reduction: *d<u>Á</u>* [ˈda] / *[ˈdɐ]
d<u>Ê</u> [ˈde] / *[ˈdɨ];

- Morphophonological processes show the asymmetry between stressed/unstressed: e.g. plurals
[fuˈniɫ] [fuˈniʃ] ‘funnel / funnels’
[ˈfasiɫ] [ˈfasɐjʃ] ‘easy / easy-PL’

  - – Pitch accent association only refers to prosodic word stress

2.1. Input Grammar: ***other properties & PW***

- Cliticization
  - 30% of phonological clitics in AS (VMF 2005)
  - Most are proclitic (Vigário 2003); 97% of produced clitics
  - Proclitics are adjoined to the level of the PW > addition of one unstressed σ
  - Enlargement of PW shapes [ɐs.tɾɐ.bɐ.ʎɐ.ˈdo.ɾɐʃ] ‘the workers (fem)

- Resyllabification
  - Not PW-bound, spans the Intonational Phrase (Frota 2000)
  - May yield a restructuring of PW (Vigário 2003) e.g. *músico* [ˈmu.si.kwɐ.fɾi.ˈkɐ.nu]/ [ˈmu.si.kɐ.fɾi.ˈkɐ.nu] ‘African musician’
  - Still, **the clustering of numerous phenomena signalling the PW in EP** > closer to Germanic Ls

2.1. Input Grammar: ***Predictions***

- The relevance of the **grammar of the target system** (i.e. the phonological system) to the development of PWs
  - EP phonology offers a rich array of evidence for the PW (though resyllabification is Romance-like: Vigário 2003)
  - Strong evidence in the input may have consequences for acquisition, as children may be aware of (at least) some of these cues
  - It is excepted that **early child speech** exhibits word-based phonology, matching the target system

2.2. Evidence for the PW in the acquisition data

PW-edges are treated differently from word-internal positions

- Word-final Coda fricatives mastered before word-internal ones (e.g.Inês) (Freitas 1997; Freitas, Miguel & Faria 2001)

festa	/ˈ*fɛʃtɐ*/	[ˈt*ɛtɐ*]	(1;9.19) ‘party’
estas	/ˈ*ɛʃtɐʃ*/	[ˈ*ɛtɐʃ*]	(1;10.29) ‘these’
versus			
bolos	/ˈ*boluʃ*/	[ˈ*boloʃ*]	(1;9.19) ‘cakes’
bonecas	/*bu*ˈ*nɛkɐʃ*/	[*mɨ*ˈ*ɲɛkɐʃ*]	(1;9.19) ‘dolls’

2.2. Evidence for the PW in the acquisition data

- Sequences of consoants to be syllabified in diff. syllables appear word-initially: sC clusters (Marta) (Fikkert & Freitas 1999; Freitas & Rodrigues 2004)

estrela	/**ʃ.ˈt***ɾelɐ*/	[**ʃˈt***ɛlɐ*]	(2;1.19) ‘star’
esticar	/**ʃ.t***iˈkaɾ*/	[**ʃ:t***iˈkaj*]	(2;2.17) ‘to stretch’

- POA assignment – word-left periphery first (Inês 1;8.2 – 1;9.19) (Costa & Freitas 2003; Costa 2004)

copo	/ˈ*kɔpu*/	[ˈ**p***atu*]	‘glass’
tampa	/ˈ*tɐ̃pɐ*/	[ˈ**p***ɐ̃tɐ*]	‘cover’
folha	/ˈfoʎɐ/	[ˈ**k**uʎɐ]	‘leaf’

2.2. Evidence for the PW in the acquisition data

- Unstressed word-initial vowels do not reduce as word-internal ones, matching the target system:

orelhas /**o**ˈɾɐʎɐʃ/ [**ɔ**lɨˈleɾɐʃ] (Inês: 1;10.29) ‘ears’

*/u/

elefante /**e**lɨˈfɐ̃tɨ/ [**i**ˈfɐ̃tɨ] (Luís: 2;0;27) ‘elephant’

*/ɨ/

orelhas /**o**ˈɾɐʎɐʃ/ [**o**ˈdɐjɐʃ] (Luís: 2;2.27) ‘ears’

*/u/

These facts confirm our expectations: word-based phonology (matching the target system) emerges early in EP.

3. A Frequency Study of Prosodic Word shape

- Language-specific frequency distributions of PW shapes in the input may constrain PW development
- Initial state with monomoraic ws > **Early words are minimally and maximally a binary foot** (e.g. Demuth & Johnson 2003)
- But variation depending on the statistical properties of the input language: emergence and development of subminimal PWs and/or PWs with more than a binary foot (e.g. Demuth & Johnson 2003, Lléo 2004, Prieto 2004)
- Analysis of PW shape frequencies in adult speech, child-directed speech and in children's early productions

3. Data

- CS: spontaneous data from 3 monolingual Portuguese children – 4.073 tokens (prosodic word forms):

João	aged 0;10.2	to	2;0.19
Inês	aged 0;11.14	to	1;10.29
Marta	aged 1;2.0	to	2;0.26

- AS: spontaneous adult speech (*Português falado 90s*)

  23.459 phonological tokens (PWs & clitics)

- CDS: spontaneous adult speech

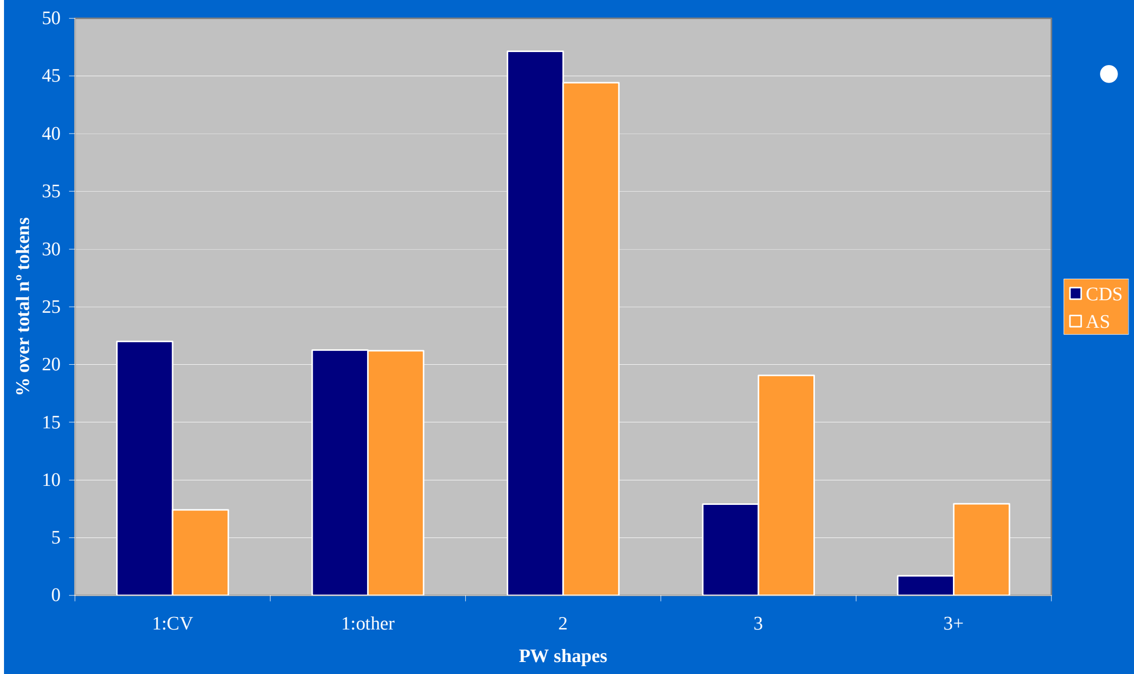
  3 first sessions of Inês (0;11.14 – 1;1.30)
  3 first sessions of Marta (1;2 – 1;4.8)
  Total of 23.207 phonological tokens (PWs & clitics)

3.1. PW shape frequencies in the input

- Child ambient language includes both CDS and AS (the 2 sorts of input have not been previously compared)
- The impact of CDS in child speech is not always clear (e.g. Ratner 1996, van de Weijer 2002)
- Some studies have concluded that CS is closer to AS (Frota & Vigário 1995)

- Breakdown of phonological tokens into PWs and clitics
  - CDS: 74.4% 25.6%
  - AS: 70.4% 29.6%
- Frequency patterns of PW shapes:
  - Monosyllabic, disyllabic
  - Trisyllabic, >3 syllables
  - Within monosyllabic: CV-shape (all PWs ending with oral V)

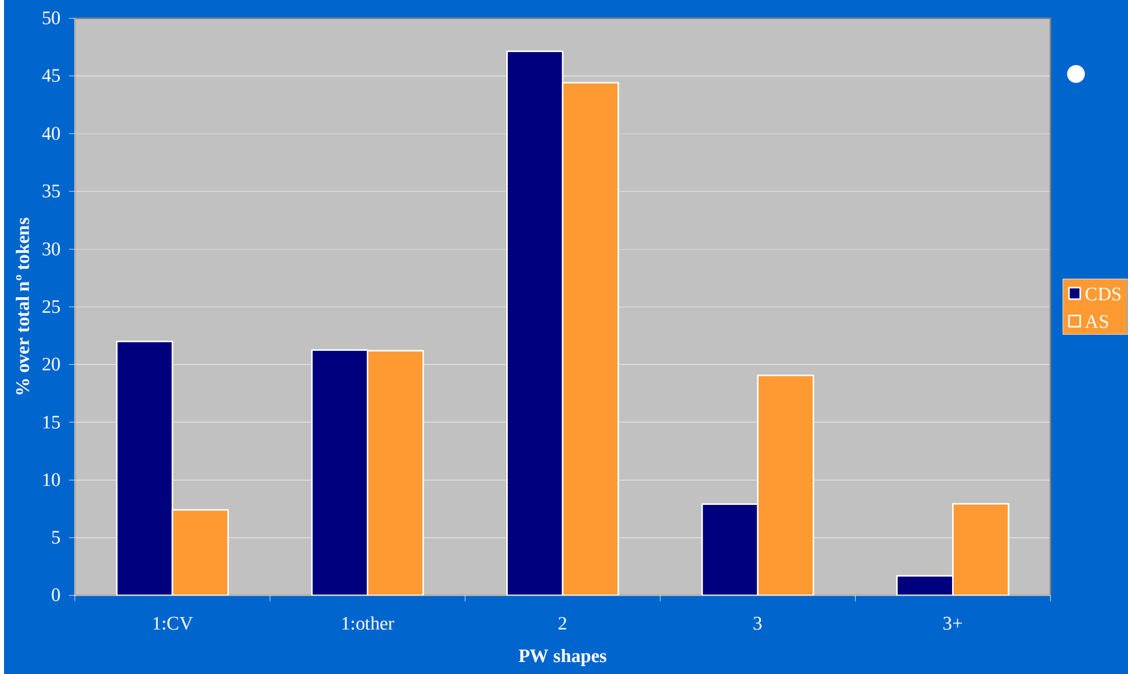
3.1. PW shape frequencies in the input



- Differences
  - CDS: monosyllabic CV shapes prevail over trisyllabic and 3+
  - AS: balanced distribution monosyllabic and PWs larger than binary foot (28.6% and 27%), monosyllabic CV and long PWs (7.4% and 8%)
- Different frequency-based predictions

- Similarities
  - Frequency of disyllabic PWs
  - Frequency of monosyllabic non-CV PWs

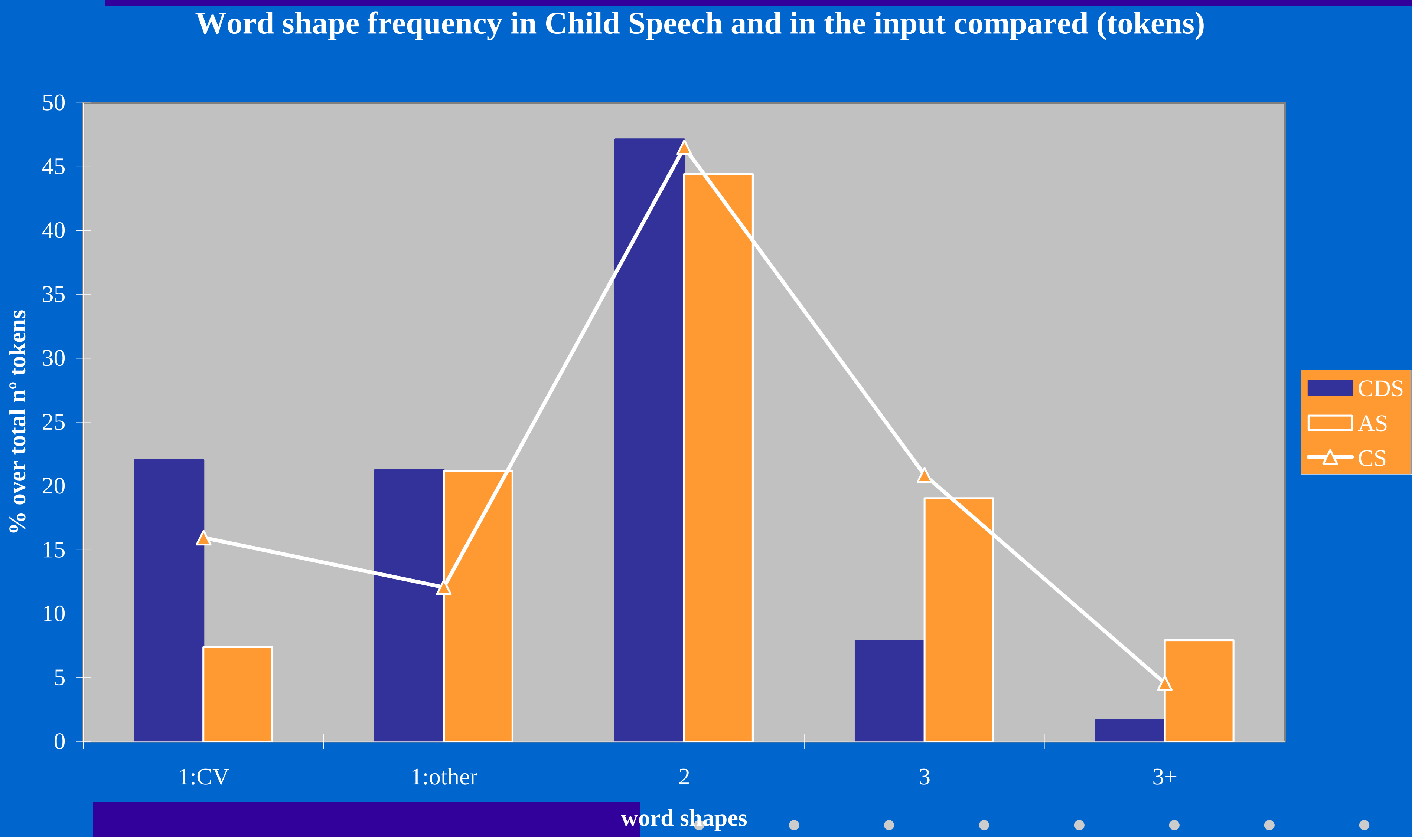
3.1. PW shape frequencies in the input



- Different frequency-based predictions
  - **CDS** : child speech will show a high incidence of subminimal PWs, while complying with maximality constraints (i.e. 3+ avoided/truncated in early speech & acquired later)

- **AS** : early child speech will show **both** the presence of subminimal PWs and larger PWs, thus not complying with constraints on word size

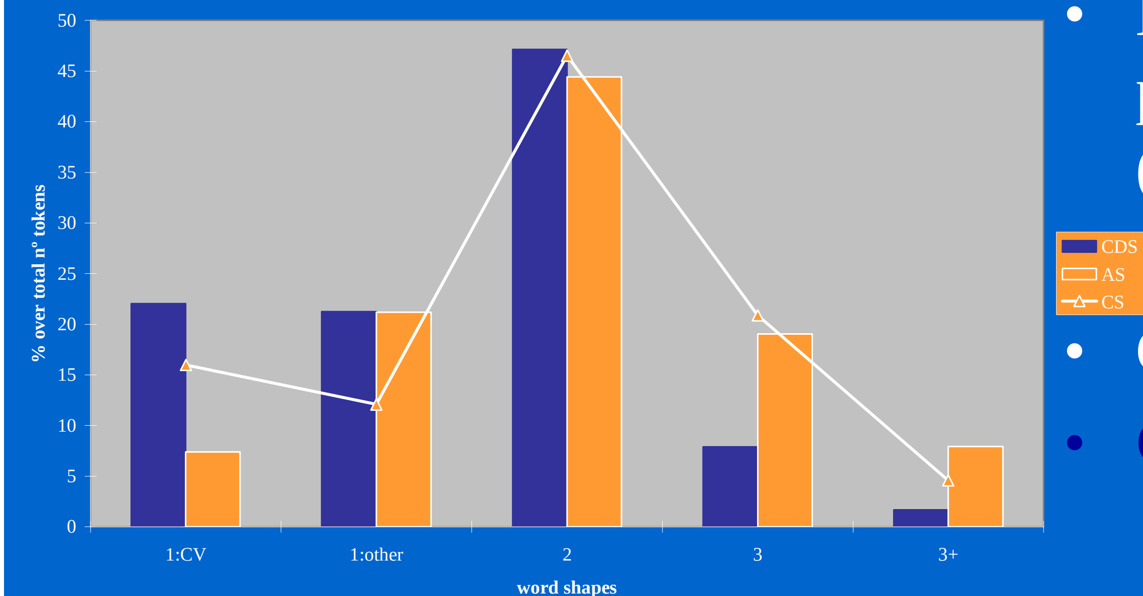
3.2. The shape of early words



The prediction based on AS frequency patterns was borne out: early CS shows BOTH subminimal and larger words

3.2. The shape of early words

Word shape frequency in Child Speech and in the input compared (tokens)



- Disyllabic shapes predominate, as expected (≈AS/CDS)
- Crucial data: 1, 3, 3+
- **CDS**: the % 1 is lower 28/**43**
  the % 3 and 3+ is much higher 25/**10**
- **AS**: the % 1 is ≈ 28/**29**
  the % 3 and 3+ is also ≈ 25/**27**

Correlation	CS	AS	CDS
CS	__	**,99***	,88
AS	**,99***	__	,91
CDS	,88	,91	__

3.2. The shape of early words: 1

- Monosyllabic CV
  - appear in large numbers
  - are produced frequently even at **later** stages: from 1;08 onwards % 1:CV/1 tokens

João 58%, Inês 59%, Marta 32%

  - **match** target word shape from the beginning (>90%)
  - No lengthening

pé /ˈpɛ/ [ˈpɛ] Inês: 1;0.25 'foot'

pé /ˈpɛ/ [ˈpe] Marta: 1;02.0'foot'

dá /ˈda/ [ˈdɐ] João: 1;02.01'give'

- Monosyllabic non-CV can be produced as CV

quer /ˈkɛɾ/ **[ˈkɛ]** Inês: 0;11.14 '(he/she) wants'

cais /ˈkajʃ/ **[ˈka]** Marta: 1;04.08 'to fall – 2nd sg'

mais /ˈmajʃ/ **[ˈmɐ]** Inês: 1;06.06 'more'

cão /ˈkɐ̃w̃/ **[ˈkɐ]** João: 1;06.18 'dog'

cais /ˈkajʃ/ **[ˈka]** João: 1;09.11 'to fall – 2nd sg'

3.2. The shape of early words: 1

- The course of development of disyllabic targets
  - truncation to 1 even in later stages
  - both iambic and trochaic targets are truncated
  - preservation of the stressed syllable

(trochaic forms more frequent in the input; the unstressed σ is reduced in both types of targets *prominence cues to PW in EP*)

- Disyllabic iambic

fugiu /fuˈʒiw/ [**ˈʃiw**] Inês: 1;01.30
‘(he/she) ran away’

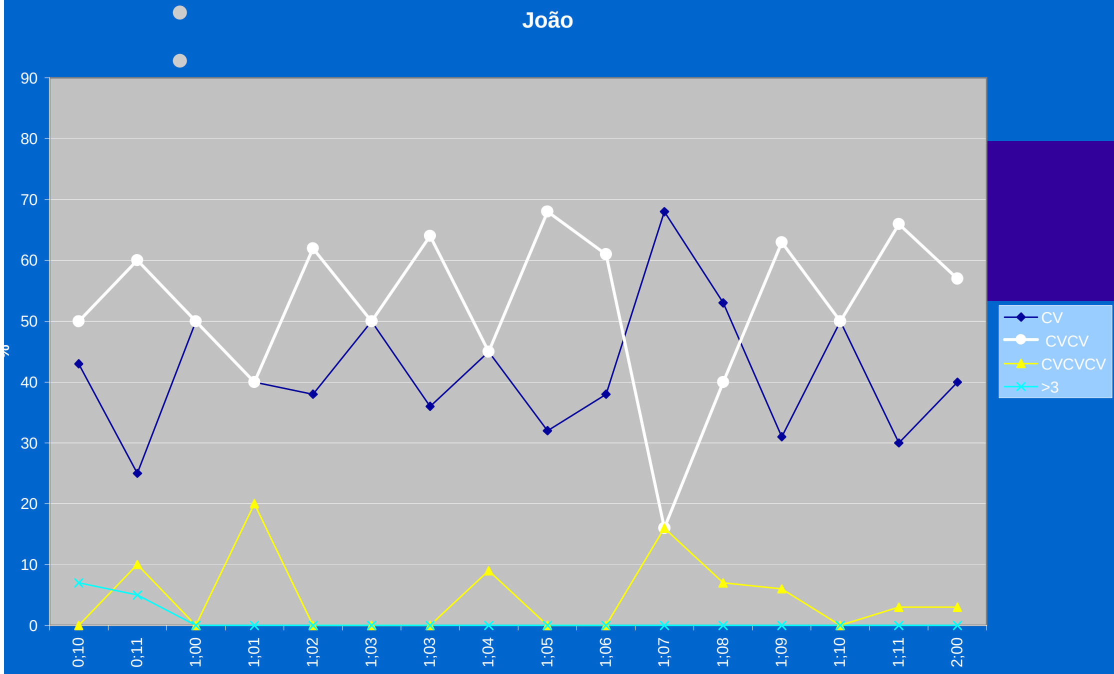
balão /bɐˈlɐ̃w̃/ [**ˈlaw**] Marta:1;03.08
‘balloon’

avô /ɐˈvo/ [**ˈbo**] João: 1;11.13
‘grandfather’

- Disyllabic trochaic

água /ˈagwɐ/ [**ˈa**] Inês: 1;01.30
‘water’

praia /ˈpɾajɐ/ [**ˈpa**] João: 2;0.19
‘beach’



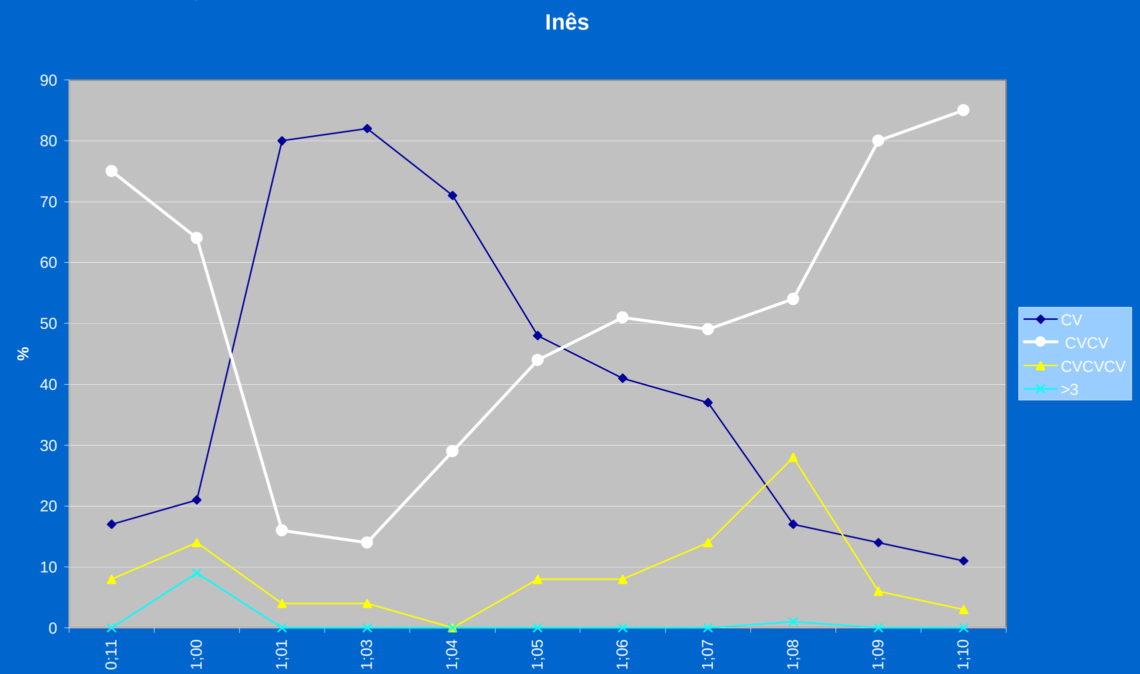
The shape of early words: disyllabic targets

Frequency of 1 shapes

Reduction of 1:other>1:CV

Truncation of 2 > 1:CV

Early words are **not** constrained by **minimality** requirements

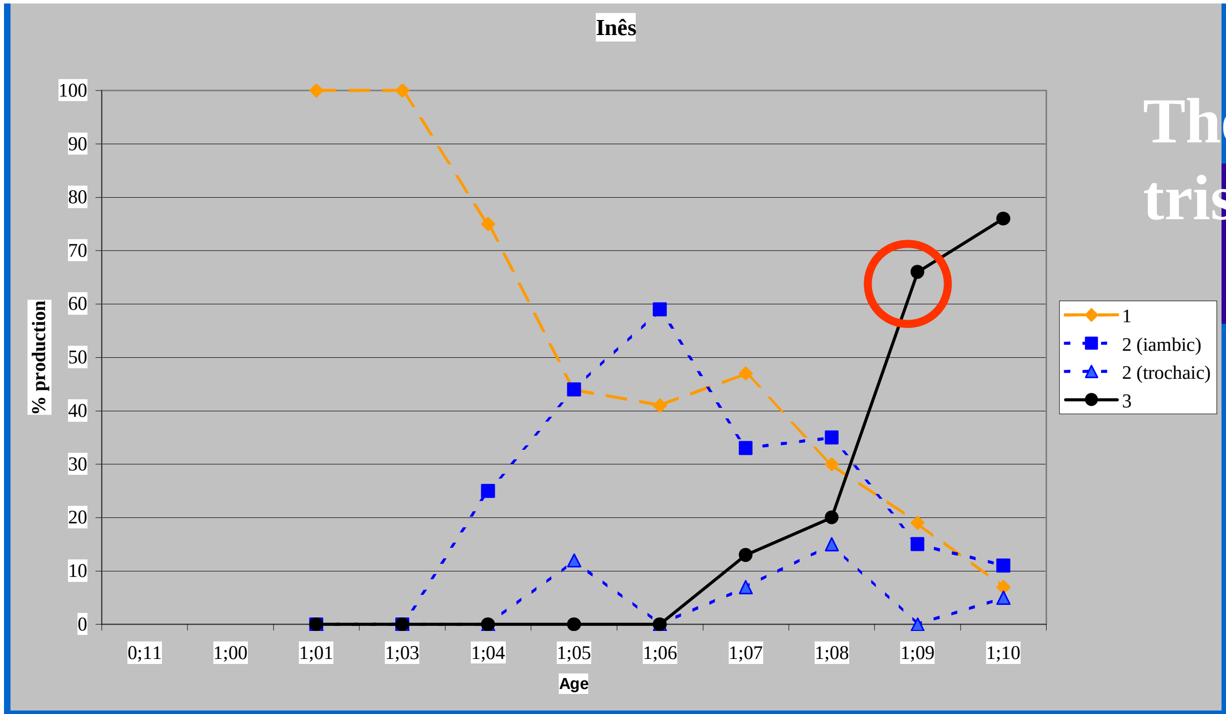


3.2. The shape of early words: >2

- Words larger than disyllabic shapes are not avoided in CS
  - Their overall frequency (25%)
  - The course of development of trisyllabic targets

- Examples

sapato /sɐˈpatu/ **[ɤɐ.ˈpa.θɐ]**
Marta: 1;02.0 ‘shoe’
banana /bɐˈnɐnɐ/ **[mɐ.ˈnɐ:.nɐ]**
Marta: 1;03.08 ‘banana’
caneta /kɐˈnetɐ/ **[kɨ.ˈle.lɐ]**
Marta: 1;05.17 ‘pen’
banana /bɐˈnɐnɐ/ **[ɐ.ˈmɐ.nɐ]**
Inês: 1;05.11 ‘banana’
sapato /sɐˈpatu/ **[pa.ˈta.tɐ]**
Inês: 1;07.02 ‘shoe’



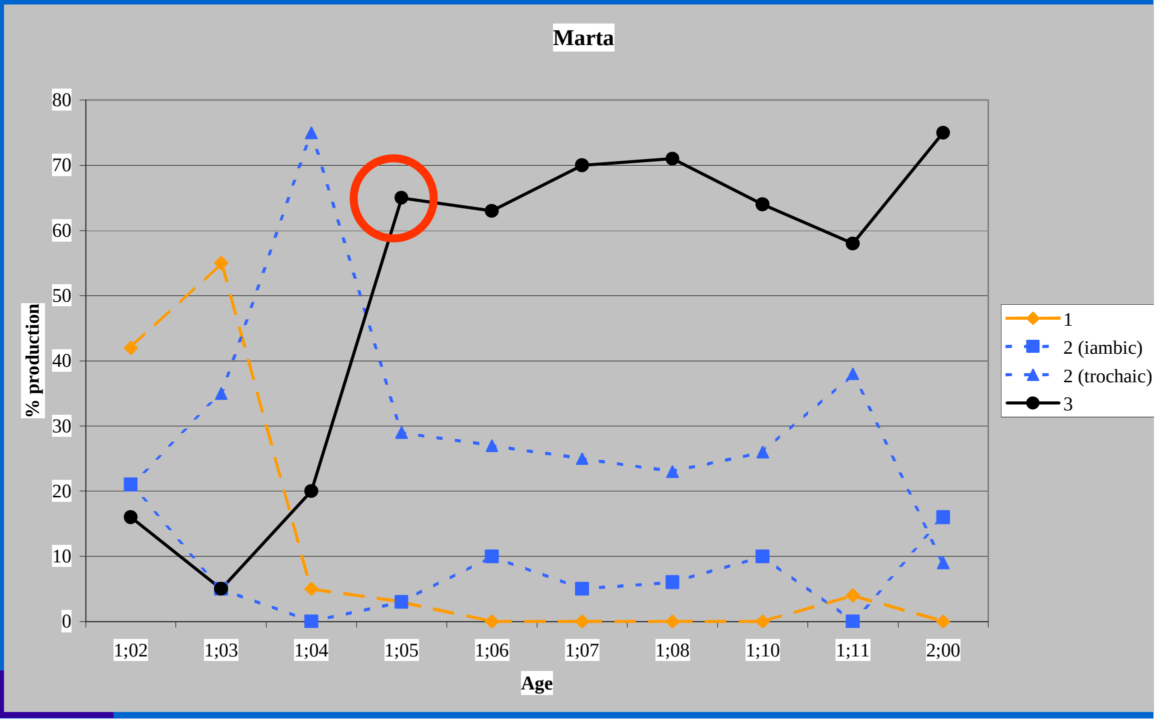
The shape of early words: trisyllabic targets

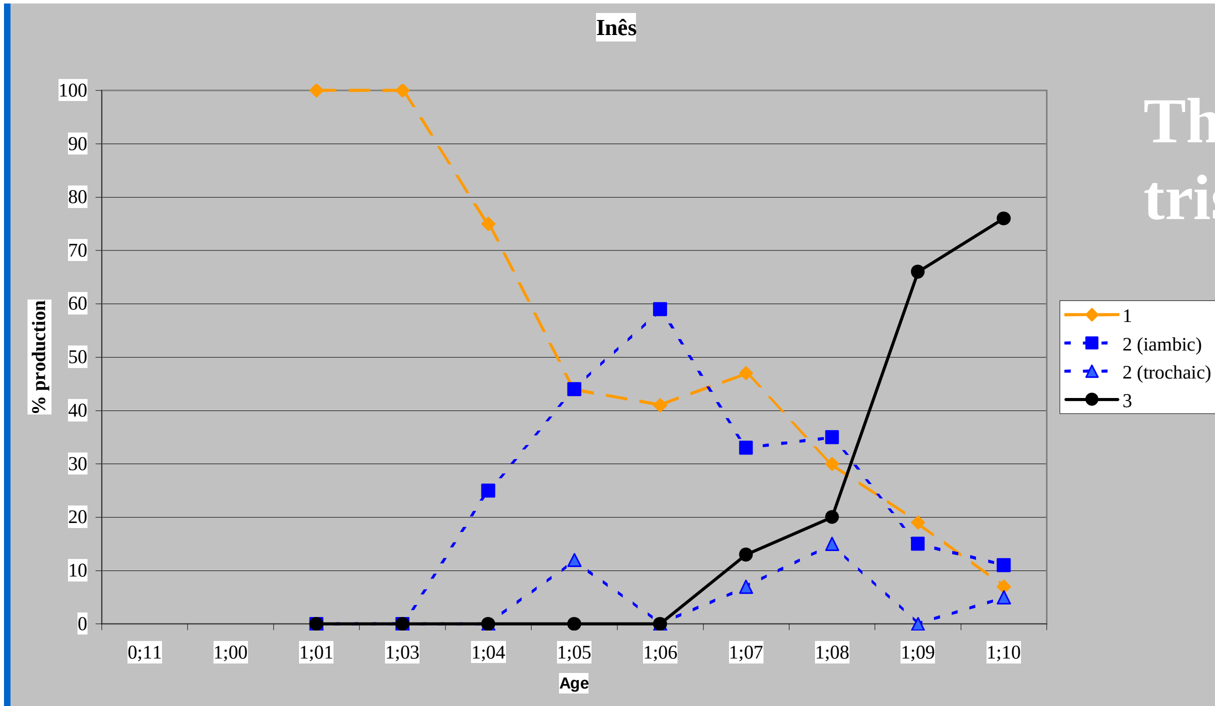
Prediction based on input frequency borne out: shapes > **binary foot** appear **early** in CS

First 3 targets appear very early: 1;01

3 word shapes emerge early: 1;02 / 1;07

Are mastered soon in the path of development





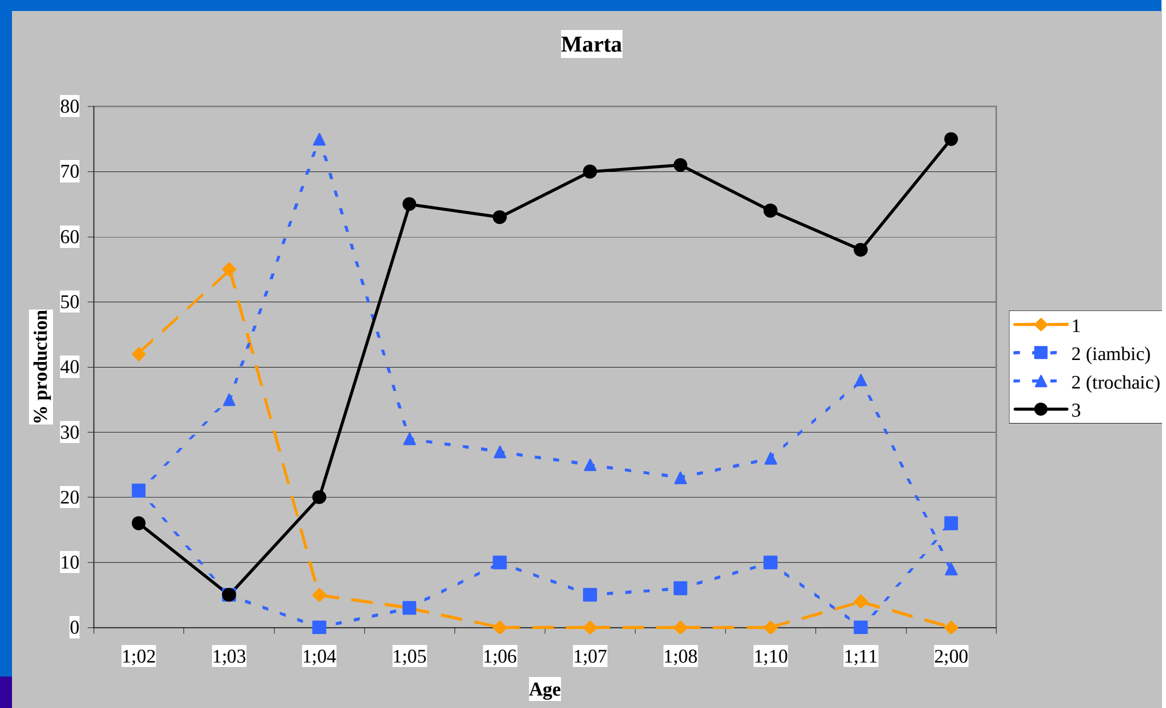
The shape of early words: trisyllabic targets

Initial stage with truncation to a 2 shape

Two strategies:

Deletion of final σ > iambic shape

Deletion of first σ > trochaic shape



Not inserted to obtain a given shape & optional

3.2. The shape of early words > 2

- About truncation to disyllabic shapes

  Preservation of the stressed σ

  Preservation of the consonant (place features) from the left-edge of PW

boneca /buˈnɛkɐ/ **[meˈɲɛ]** 1;05.11 'doll'

mamoca /mɐˈmɔkɐ/ **[mɐˈmɔ]** 1;05.11 'little breast'

morangos /muˈɾɐ̃guʃ/ [ˈ**m**ɐ̃guʃ] 1;05.17 'strawberrys'

querido /kɨˈɾidu/ **[ˈkidu]** 1;05.17 'dear'

- Prosodic fillers
  - Initial σ added to material that realizes the target PW
  - Regardless of PW category
  - Regardless of PW size

pato /ˈpatu/ [ɐ.ˈtɐ]/[ˈtɐ] João: 0;11.06 'duck'

dá /ˈda/ [ɐ.ˈda]/[ˈda] Inês: 1;0.25 'give (me)'

mola /ˈmɔlɐ/ [ɨ.mɔˈ.lɐ]/[ˈmɔ.lɐ] Marta: 1;05.17 'spring'

Early words in EP are NOT constrained by minimality or maximality requirements

4. Summary and Discussion

- Summary of findings
  - PW shape frequencies in the input contribute to explain PW acquisition (AS in particular)
  - As predicted by the frequency patterns, 3 targets appear early (1;01/1;02) & are produced early (1;02/1;07); 1:CV remain frequent until later stages
  - The properties of the input grammar concur to promote the same effect

- Discussion
  - In EP, both the grammar and frequency effects promote the early production of the ≠ word shapes
  - Strong evidence for PW in the input grammar may strengthen the frequency effects

  e.g. PW-edges are relatively well-delimited (> other RLs) a tendency is expected to faithfuly reproduce edges in CS

4. Discussion: Grammar and Frequency

- Examples of possible interactions
  - 3 PWs are acquired early
  - But initial stage of truncation Two of the strategies children display involve preservation of the PW left-edge, the most prominent PW-edge in EP grammar
  - 2 PWs may show truncation to 1, even at the later stages. Both iambic and trochaic targets are reduced to the stressed syllable, in line with prominence-related cues to the PW (i.e. salience of word stress, unstressed V reduction)
  - Main direction of cliticization in EP (proclisis) increments the frequency of larger words, adding to evidence vs. a maximality constraint

Resyllabification makes evidence for PW less strong then in GLs

• **Still, little data available!**

4. Discussion: Grammar and Frequency

- English, Spanish, Catalan, EP
  - Grammar: Eng > EP > Sp, Cat

  Prediction: Early production of the ≠ word shapes

  But shapes larger than a binary foot Sp, EP > Eng, Cat

(Roak & Demuth 2000, Demuth & Johnson 2003, Lléo & Demuth 1999, Lléo 2004, Prieto 2004)

  - Frequency: Sp (≈30%), EP (27%) > Cat (15%) > Eng (≈5%)

  But Sp, EP > Eng, Cat   If frequency alone explains the early appearance in Sp and EP, it does not explain the fact that they seem to emerge equally late in Cat and Eng

  ✓ If a Grammar & Frequency interaction is assumed: a considerably higher frequency in Cat, but much strong grammar cues in Eng.

Obrigada !

This work is part of the research developed within the SILC project, held by Onset-CEL (Univ. Lisboa), in collaboration with Univ. Minho. It will appear in *Language and Speech* Special Issue on Prosodic Words

Special thanks: the children and their mothers; Ana Lúcia Santos, Carla Soares; K. Demuth and two *Language and Speech* Reviewers; FLUL for financial support.

2.1. Input Grammar: ***Predictions***

- The relevance of the **grammar of the target system** (i.e. the phonological system) to the development of PWs
  - EP children should exhibit **early** development of ω
  - We should expect evidence for ω in EP to emerge earlier than in other Romance languages
  - Depending on the weight and/or frequency of the various grammatical cues in the input
    - EP children may pattern like Germanic children (if EP Romance-like **resyllabification** is not that important/salient)
    - Or show an **intermediate** speed of development between Germanic and Romance languages

Lower bounds on early ωs: *No minimality*

- Word shape of children's productions shows
  - ➢ High percentage of monosyllabic shapes (monomoraic)
- The high incidence of subminimal words is **NOT** consistent with a word shape frequency effect
  4% of monosyllabic words in the input;
  6% of monomoraic words in adult most frequent words
- Hypothesis: the result of a conflict between patterns with rightmost prominence, given by higher-level prosody (ϕ, I), and pattens with leftmost prominence (word-level stress): CV C'V
  C'V CV
  monosyllabic

Summary

- EP children display **early** development of ω, as shown by the difference between ω-edges and word-internal positions (consistent with the properties of the input grammar). A comparison with Germanic Ls and other Romance Ls would enligthen the issue of the frequency vs. weight of phonological cues.
- ***No maximality***: early production of trisyllabic words (consistent with both a frequency and a grammar effect)
- ***No minimality***: high incidence of subminimal words, against the low frequency in the input (a strategy to deal with a prominence conflict)
- **No link** between maximality and the early emergence of protomorphemes (an alternative view: the prosodic filler hypothesis)