

*Workshop “Domínios: Prosódia e Sintaxe”*

Campinas – 16 e 17 de Abril, 2007

*Frequência de unidades e padrões fonológicos no  
Português Europeu e no Português do Brasil –  
uma prospectiva*

Marina Vigário

Sónia Frota

Fernando Martins

Universidade de Lisboa – Laboratório de Fonética / Onset - CEL

POCTI-SFA-17-745

- 
- 
- 

## Estrutura da apresentação

1. Enquadramento muito breve
2. Resumo das funcionalidades do *FreP*
3. Demonstração
4. Aplicações anteriores do *FreP*
5. Comparação PE/PB - perspectivas futuras

•  
•  
•

# 1. Enquadramento: a importância da informação de frequência na análise linguística (exemplos concretos)

- Fonologia (Geral)

- Palavras e suas combinações frequentes >> redução

Inglês: Jurafsky, Bell e Girand 2002,  
Neerlandês: Booij 1995; PE: Vigário 2003

- Formações irregulares freqs >> não regularização

Bybee e Hopper 2001

- Ativação de restrições, e.g. palavras monossílabas ↓ Fre >> Palavra Prosódica Mínima

Italiano: Thornton 1996; Peperkamp 1997  
PE: Vigário 2003 vs. Booij 2004

- Aquisição da Fonologia

- ↑ Fre de Codas preenchidas >> velocidade de aquisição

Neerlandês vs. PE, Fikkert e Freitas 1999

- ↑ Fre *ω*s monomoraicas >> produção de *ω*s submínimas
- Francês e Catalão vs. Inglês e Neerlandês,  
Demuth & Johnson 2003, Prieto 2006

- ↑ Fre padrões fonológicos no léxico >> produção de sequências segmentais

Beckman e Edwards 2000

...

- 
- 
- 

## 1. Enquadramento: Frequências na ‘Fonologia’ do PE

- **Português Fundamental**
  - tipos silábicos *Andrade e Viana 1994, Vigário e Falé 1994, Viana et al. 1996*
  - segmentos fonéticos *Viana et al. 1996*
  - distribuição do acento (palavra morfosintáctica) *Viana et al. 1996*
  - palavras morfosintácticas e redução fonética *Vigário 2003*
- **Outros**
  - segmentos fonéticos *Delgado Martins 1975*
- **Sistemas de síntese (ferramentas integradas)** *Viana, c.p.; Almeida, c.p.*

Ferramenta para análise de Frequências do domínio público?  
Aplicável a difs. corpora e difs. unidades (morfo)fonológicas ?

- 
- 
- 

## 1. Enquadramento – o *FreP*

Criação de uma nova ferramenta para extracção de informação de frequência no Português



***FreP*** 

- 
- 
- 

## 2. O *FreP* – breve descrição

- Extrai informação sobre unidades e padrões **fonológicos** com base em **texto escrito**, de acordo com as **convenções ortográficas** estabelecidas.
- Fonologia lexical – inclui processos fonológicos obrigatórios; exclui os opcionais
- Corre sobre ficheiros de texto comuns, não-formatados (e.g. txt)



## 2. O *FreP* – breve descrição da 1ª versão (2004-2005)

- Localização do Acento (palavras agudas, graves e esdrúxulas)
-  Identificação e contagem de PW
- Identificação e contagem de Clíticos
- Identificação e contagem de Proclíticos e Enclíticos
-  Identificação e contagem de Vogais
- N° Sílabas por Palavra
-  Tamanho de PW (em n° de  $\sigma$ ; > n° segmentos)
-  Tamanho de Clíticos (em n° de  $\sigma$ ; > n° segmentos)

•  
•  
•

## 2. O *FreP* – novas funcionalidades da 2ª fase (2005/06)

- Delimitação das sílabas
- ↓
- Identificação dos tipos silábicos (CV, V, CVC, ...)
  - em função da posição na palavra  
(inicial, medial, final)
  - em função da presença / ausência de acento
- Identificação das grandes classes de segmentos  
(C, V, G)
- Determinação da distribuição do acento na palavra  
(final, penúltimo, antepenúltimo, inicial)

+ Forte investimento na componente visual da ferramenta

• • • • • • • • • •

- 
- 
- 

## 4. O *FreP* – novidades da versão 0.9 Beta 1

- Localização do acento em função do número de sílabas da palavra (agudas, graves e esdrúxulas em palavras dissilábicas...)
- Identificação e contagem de segmentos / transcrição fonética
  - 1ª fase concluída: todas as consoantes e glides
- Identificação e contagem dos traços de ponto de articulação consonântico (Labial, Coronal, Dorsal)
- Listagem ordenada da frequência de palavras únicas

- 
- 
- 

## 4. O *FreP* – Mais informação

- Página electrónica – <http://www.fl.ul.pt/LaboratorioFonetica/FreP>
- Manual – <http://www.fl.ul.pt/LaboratorioFonetica/FreP>
- Distribuição limitada da ferramenta (investigação; teste)

Algumas características adicionais:

- Linguagem de programação – C++
- Sistema de janelas/comandos baseado no formato Windows + ficheiro Excel
- Domínio público: disponibilização à comunidade científica

- 
- 
- 

## 4. O *FreP* – novas funcionalidades

*Demonstração*

*FreP* 

- 
- 
- 

### 3. Aplicações anteriores do *FreP*

#### Áreas de aplicação na análise linguística

- Restrição de *palavra mínima* no Português<sup>1</sup>
- Aquisição dos formatos de palavras<sup>2</sup>
- Frequência vs. gramática na distribuição dos pronomes verbais clíticos<sup>1</sup>

<sup>1</sup> Vigário, Martins & Frota (2004/2005) - *APL'04 / Actas'05*

<sup>2</sup> Vigário, Freitas & Frota (2004/2006) – *Second Lisbon Meeting on Child Language / Language & Speech*

- 
- 
- 

### 3. Aplicações anteriores do *FreP*

- Frequências de tipos silábicos na fala adulta<sup>3</sup>
  - tipos silábicos por posição
  - tipos silábicos X acento de palavra
- Aquisição dos tipos silábicos<sup>3</sup>
- Classes de segmentos<sup>3</sup>
- Aquisição do ponto de articulação consonântico<sup>4</sup>

3. Frota, Freitas & Vigário (2006) – *IASCL 2006*; Vigário, Martins & Frota (2006); Freitas, Costa, Frota, Martins & Vigário (2006) - *Selected Papers APL'06* ; Frota, Vigário & Martins (2006) - *LREC2006*

4. Costa, Freitas, Frota, Martins & Vigário (2006) - *APL2006*

- 
- 
- 
- 
- 
- 
- 
-

- 
- 
- 

### 3. Aplicações anteriores do *FreP*

#### Terapia da Fala

- Uso na construção de procedimentos de avaliação fonológica a crianças entre os 3 e os 6 anos  
(Guerreiro, em curso; Silva, em curso)
- Estudo do léxico de crianças (2 aos 5 anos)  
(Mina, em curso)
- Correlatos fonético-fonológicos das emoções na fala (e.g. velocidade de elocução - nº sílabas/seg)  
(Correia, em curso)

•  
•  
•

#### 4. *FreP* aplicado à comparação PE/PB

### Comparação Português Europeu / Português Brasileiro

Adaptações necessárias (colaboração de Flaviane Fernandes em projecto submetido à FCT)

Pequeno teste:

Corpus Português Oral. Documentos Autênticos CD-ROM  
(CLUL/Instituto Camões)

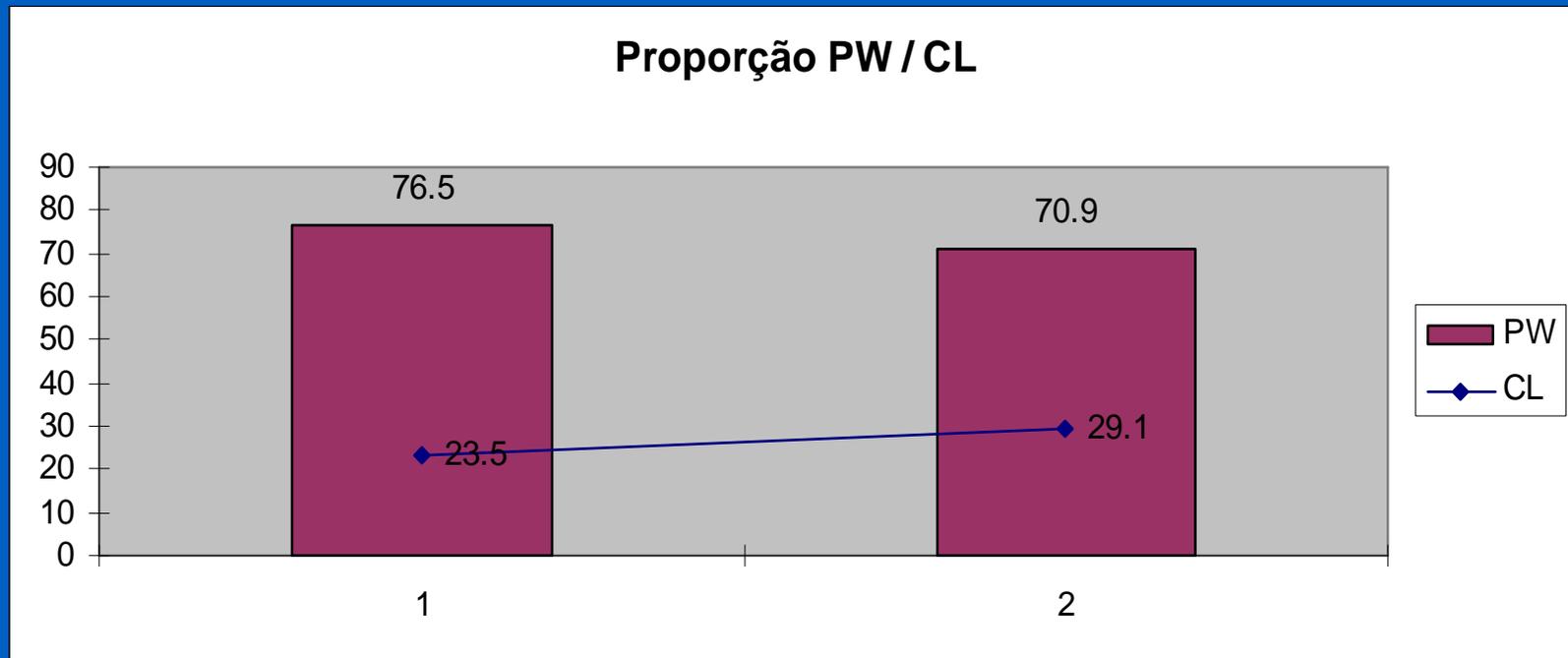
→ entrevistas orais; duas variedades > dados comparáveis

→ intervenção muito simples nos *corpora* para resolver problemas particulares (algumas abreviaturas; consoantes mudas)

• • • • • • • • • •

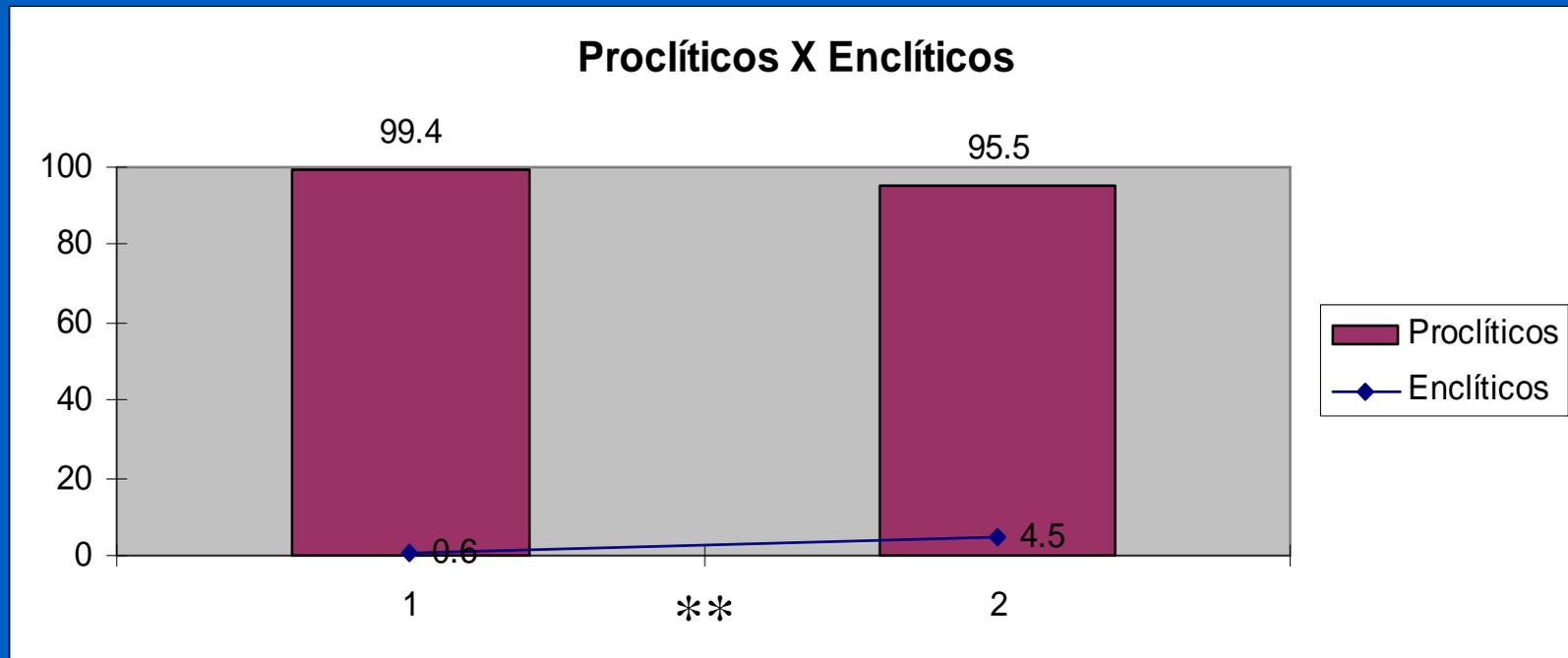
- 
- 
- 

## 4. Comparação PB/PE – Proporção PW / CL



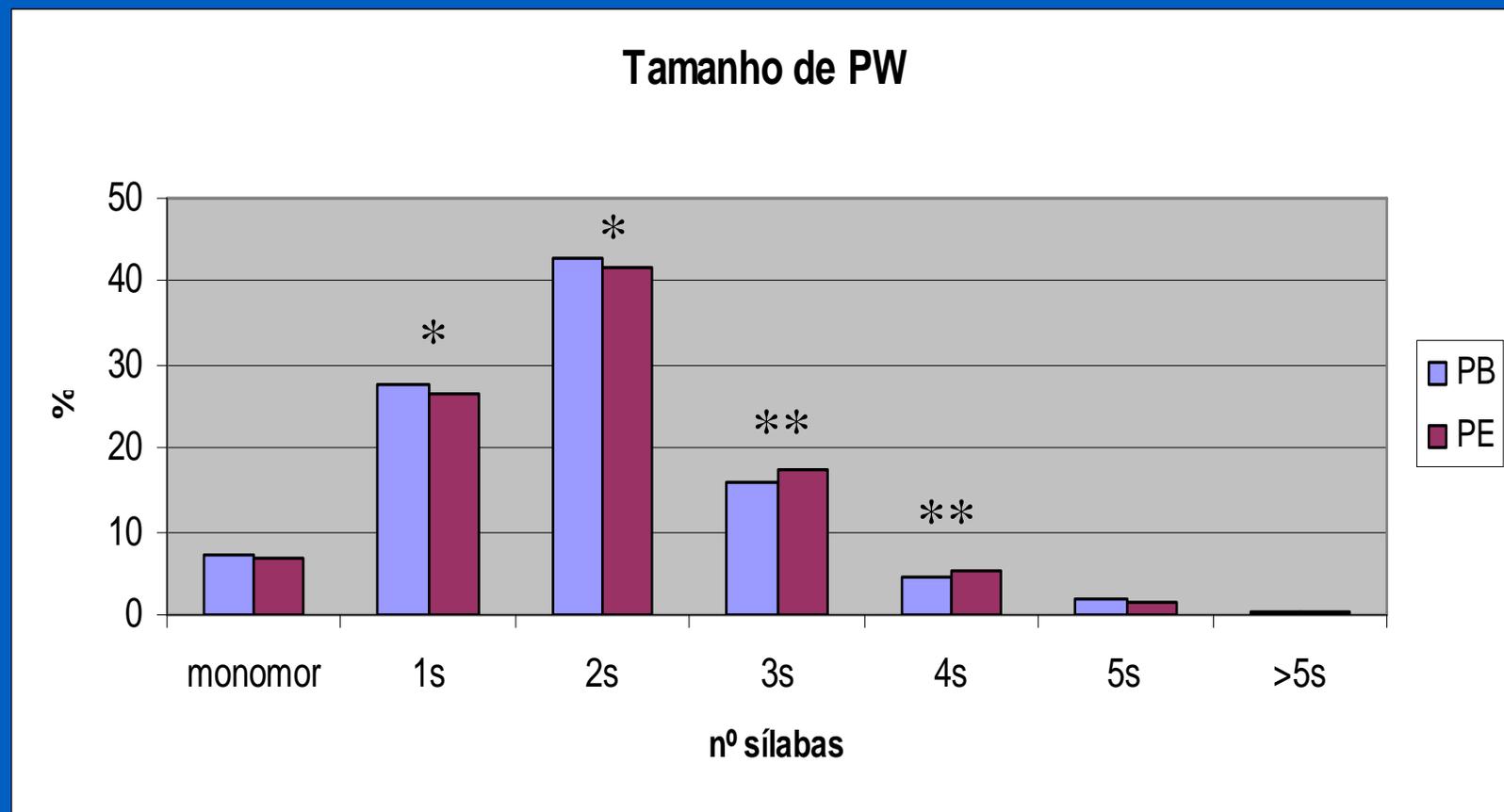
- 
- 
- 

## 4. Comparação PB/PE – Proporção Enclíticos/Proclíticos



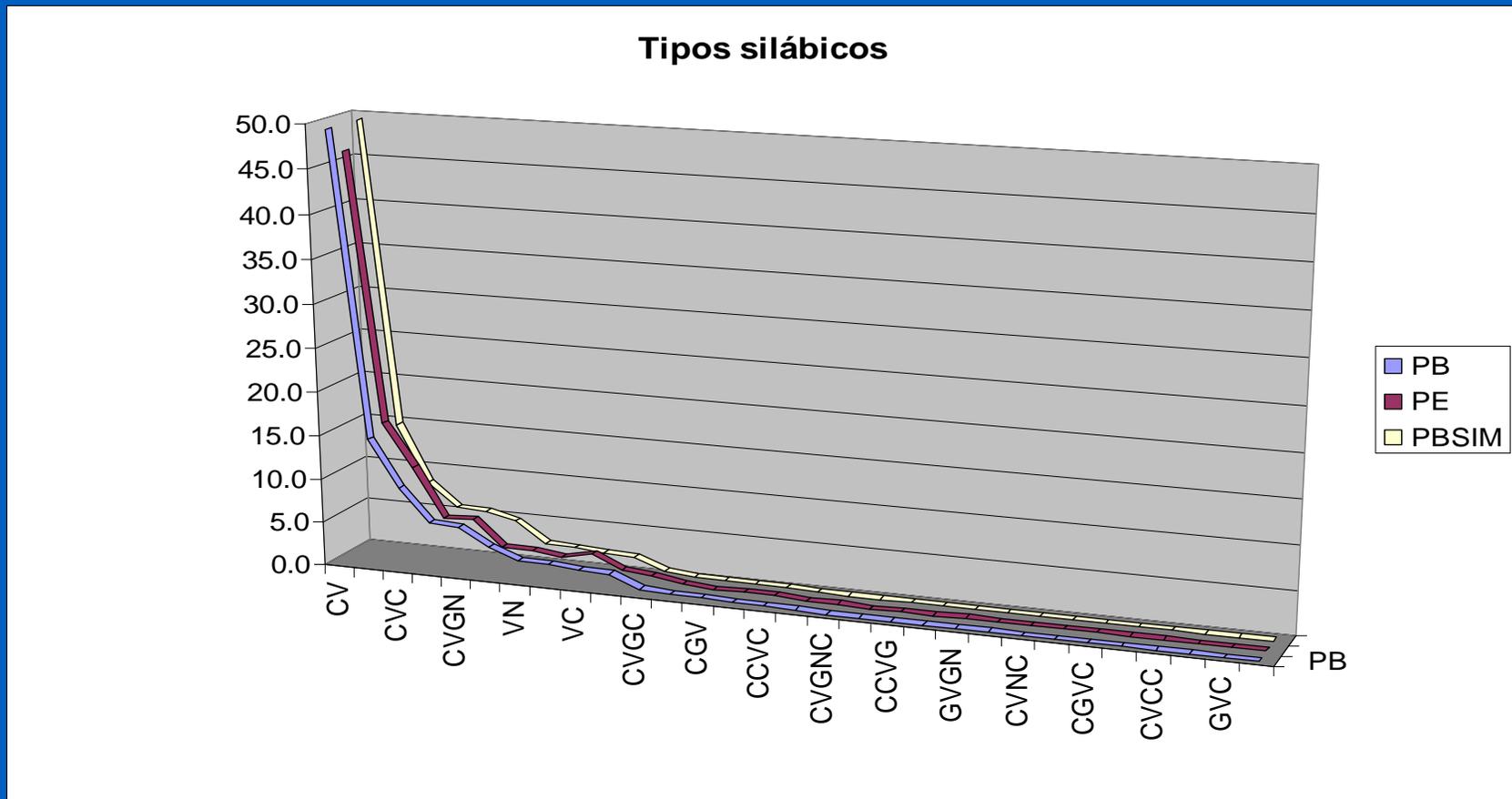
pronomes átonos: v

## 4. Comparação PB/PE – Tamanho de PW



tamanho: v

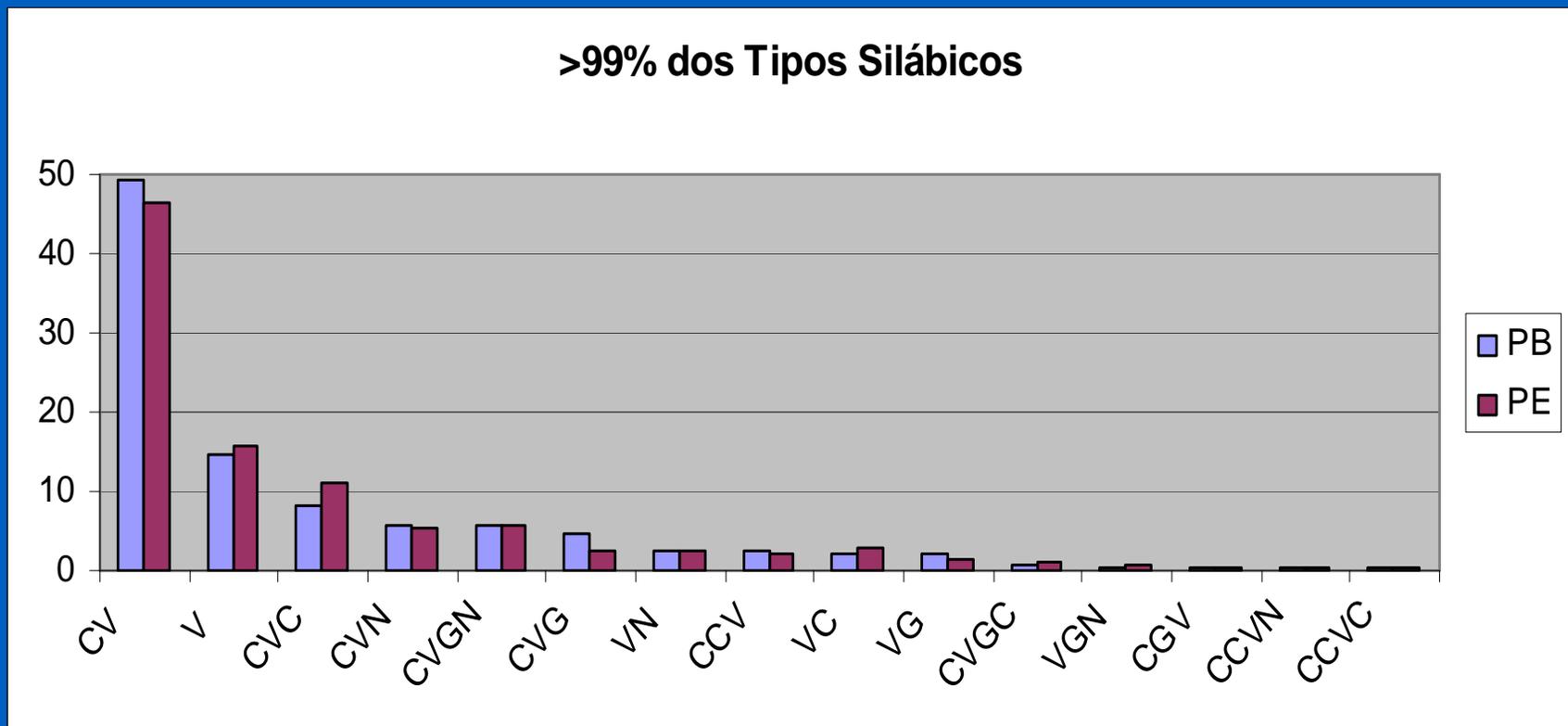
## 4. Comparação PB/PE – Tipos silábicos



$$\text{PBSIM} = \text{CVG} + \text{CVL} / \text{CVC} - \text{CVL}$$

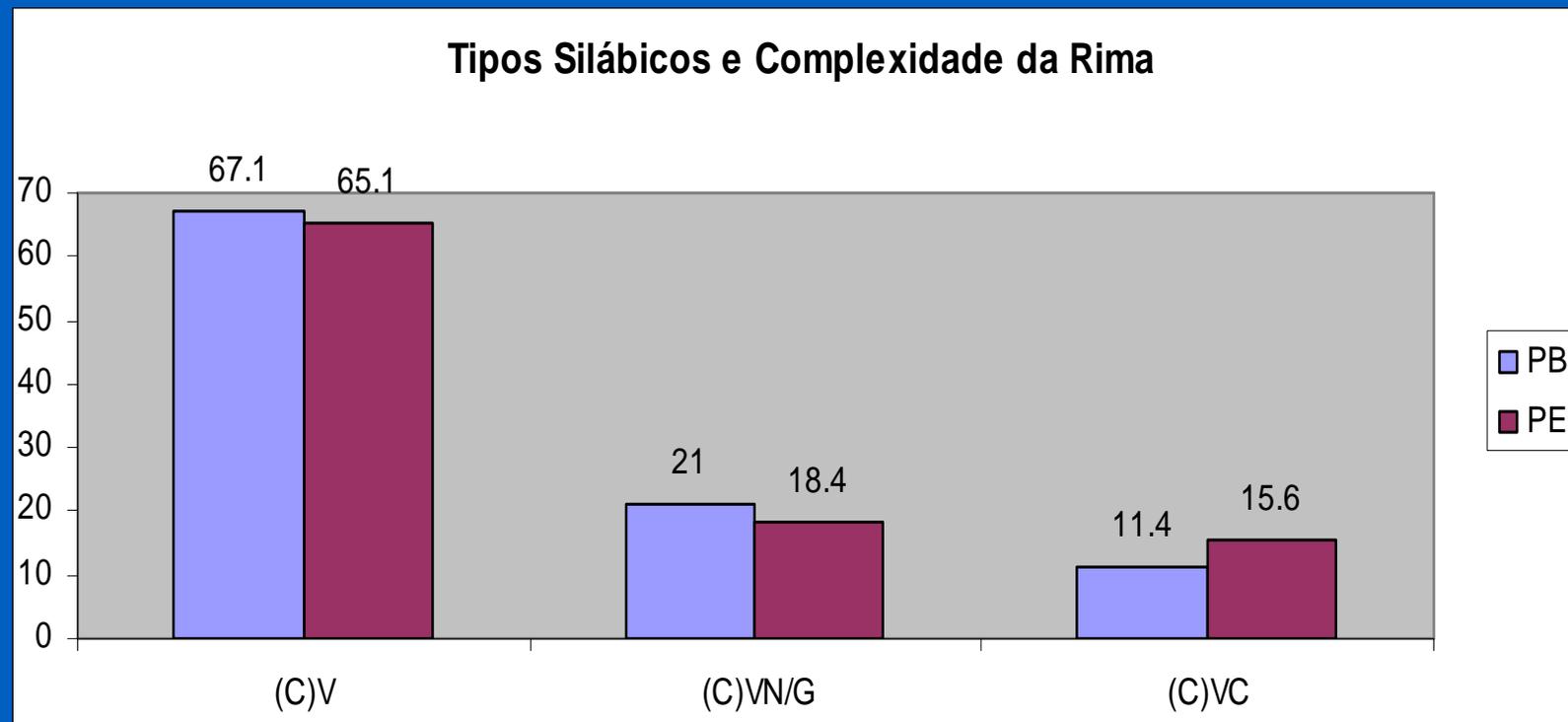
- 
- 
- 

## 4. Comparação PB/PE – 30→15 Tipos Silábicos (99.1% PB / 99.5% PE)



**PB** >CV: padrão alternante; <V: omissão de V inicial;  
<CVC: Complexidade silábica;

## 4. Comparação PB/PE – Tipos de Rima

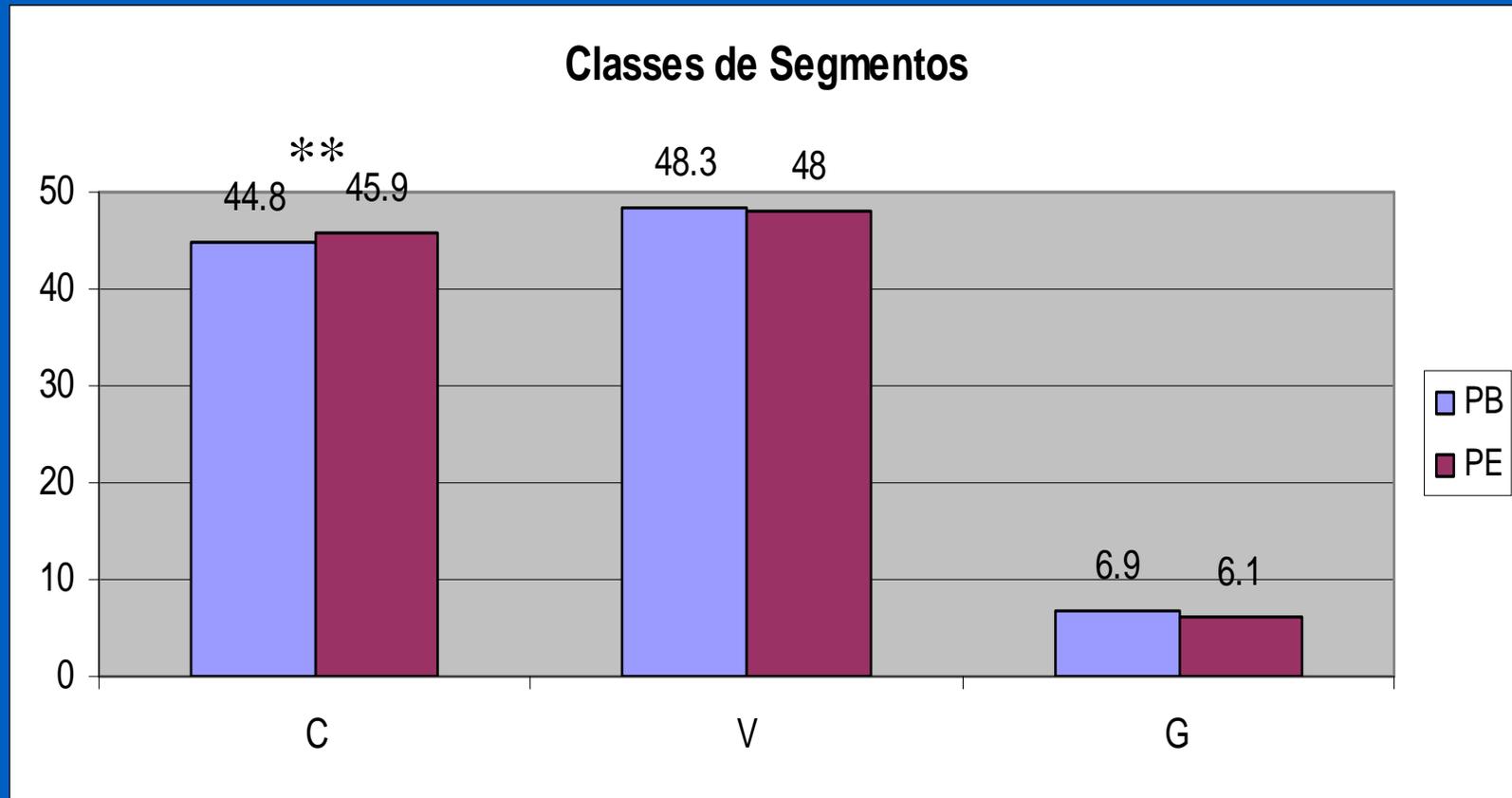


PE: >CVC; < (C)V/VN

(C)VC: v

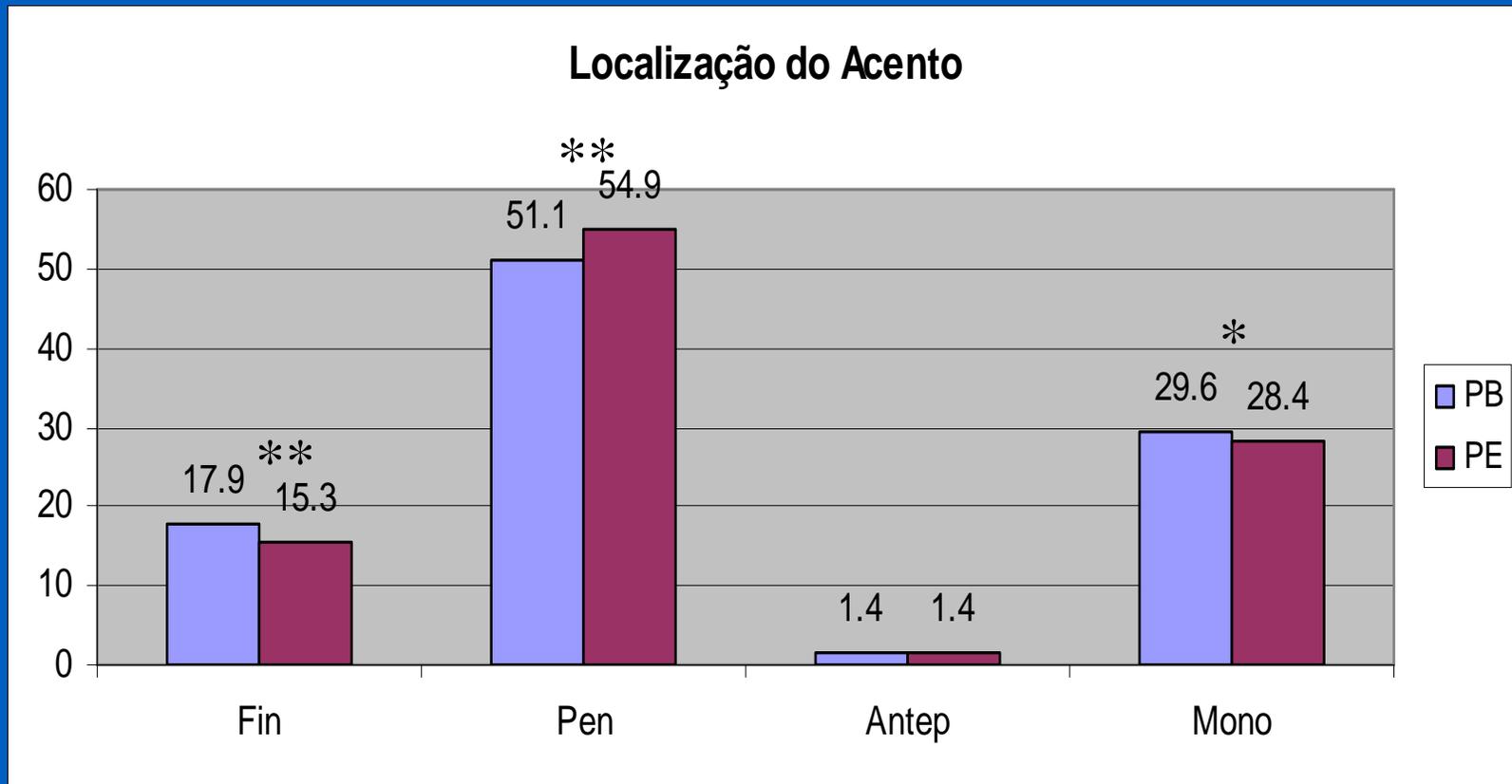
- 
- 
- 

## 4. Comparação PB/PE – Classes de Segmentos



C e V: v

## 4. Comparação PB/PE – Localização do Acento



•  
•  
•

## 6. Perspectivas futuras: trabalho a expandir

Expansão, adaptação e correcção do FreP	Aferir a importância dos factores:	Comparar dados do adulto e dados de aquisição
Constituir base de dados de referência	<i>texto, instrução, profissão, idade, sexo, estilo...</i>	<b><i>Extraír consequências para a análise gramatical</i></b>
Medir a fiabilidade	Comparar Variedades/Épocas do Português	Outras aplicações: Terapia da fala Ensino Invest. Criminal ....

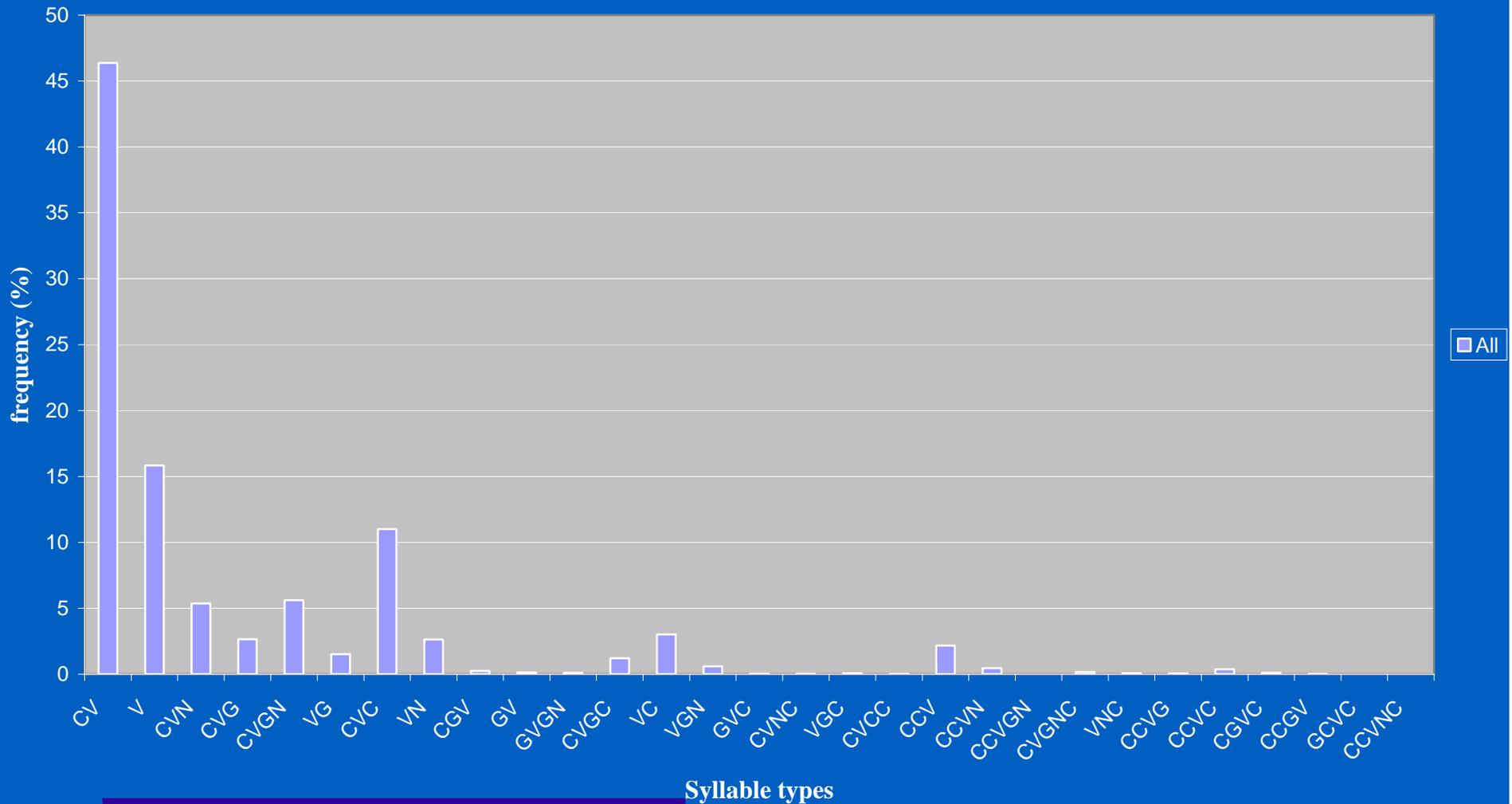


*Obrigado!*



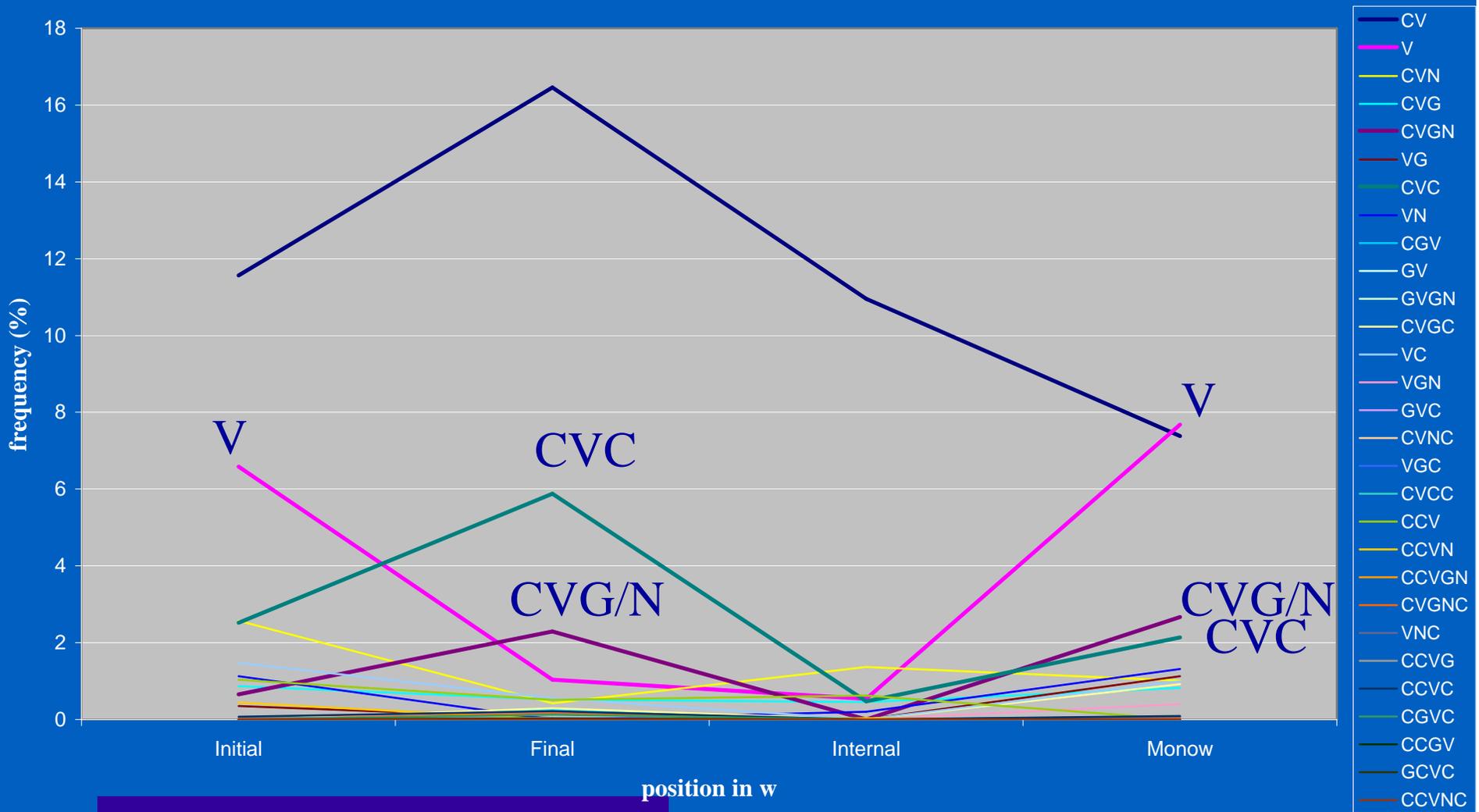
## 5. Resultados novos (1): Frequências de tipos silábicos na fala adulta

ADS



# 5. Resultados novos (1): tipos silábicos por posição

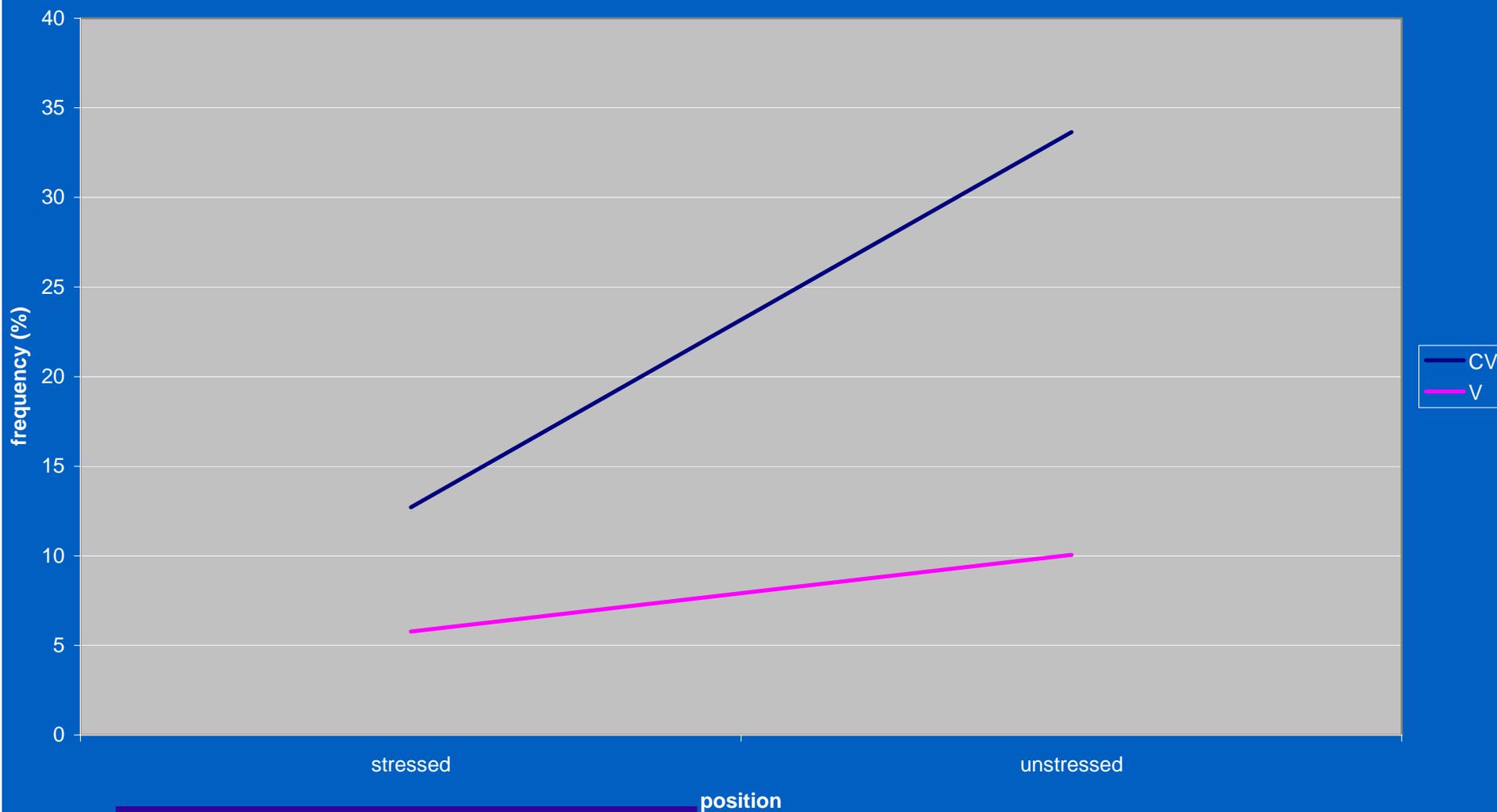
ADS - Syllables by position in w



- 
- 
- 

## 5. Resultados novos (1): tipos silábicos X acento de palavra - CV, V

ADS - The effect of stress



# 5. Resultados novos (1): tipos silábicos X acento de palavra (restantes tipos)

ADS - The effect of stress

