

# ACENTO NO PB: ANÁLISE QUANTITATIVA

# O ACENTO

<b>CAT-ACENTUAL</b>	<b>% DE TYPES</b>	<b>% DE TOKENS</b>
Monossílabas	2,39%	32,16%
Oxítonas	30,21%	27,31%
Paroxítonas	63,46%	38,89%
Proparoxítonas	3,90%	1,62%
4-sílabas	0,04%	0,03%
<b>Total</b>	<b>100,00%</b>	<b>100,00%</b>

# O ACENTO - Teorias

EST. SIL. por CAT.ACENT.	% de Tipos	Frequência de Tipos
<b>oxítona</b>	<b>100%</b>	<b>5190</b>
&CV-CV-CVG*	27,53%	1429
&CV-CVG*	25,09%	1302
&CV-CVC*	21,37%	1109
&CV-CV-CVC*	13,56%	704
&CV-CV-CV-CVG*	12,45%	646
<b>paroxítona</b>	<b>100%</b>	<b>13776</b>
&CV-CV-CV*	34,73%	4784
&CV-CV*	24,35%	3355
&CV-CV-CV-CV*	21,44%	2953
&V-CV-CV-CV*	10,84%	1494
&V-CV-CV*	8,64%	1190
<b>proparoxítona</b>	<b>100%</b>	<b>893</b>
&CV-CV-CV*	33,48%	299
&CV-CV-CV-CV*	23,29%	208
&CV-CV-CV-CVS*	16,35%	146
&CV-CV-CVS*	15,12%	135
&V-CV-CV*	11,76%	105
<b>Grand Total</b>		<b>19859</b>

# O ACENTO - Teorias

EST. SIL. por CAT.ACENT.	% de Tipos	Frequência de Tipos	
<b>oxítona</b>	<b>100%</b>	<b>5190</b>	
&CV-CV-CVG*	27,53%	1429	
&CV-CVG*	25,09%	1302	
&CV-CVC*	21,37%	1109	28662
&CV-CV-CVC*	13,56%	704	
&CV-CV-CV-CVG*	12,45%	646	
<b>paroxítona</b>	<b>100%</b>	<b>13776</b>	
&CV-CV-CV*	34,73%	4784	
&CV-CV*	24,35%	3355	60205
&CV-CV-CV-CV*	21,44%	2953	
&V-CV-CV-CV*	10,84%	1494	
&V-CV-CV*	8,64%	1190	
<b>proparoxítona</b>	<b>100%</b>	<b>893</b>	
&CV-CV-CV*	33,48%	299	3701
&CV-CV-CV-CV*	23,29%	208	
&CV-CV-CV-CVS*	16,35%	146	
&CV-CV-CVS*	15,12%	135	
&V-CV-CV*	11,76%	105	94872
<b>Grand Total</b>		<b>19859</b>	

# A questão central da pesquisa

- Ambas as teorias usam a marcação lexical para dar conta da totalidade do português.
- Se eliminarmos a marcação (arbitrária) lexical, é possível criar alguma abordagem bem sucedida para a questão do acento?
- Ou seja, é possível criar um modelo que descreva o comportamento do acento sem usar a marcação arbitrária?

# Corpus ABG - montagem

- Corpus Oral

- C-oral do Brasil (UFMG)
- Iboruna (UNESP – Rio Preto)
- Projeto SP (USP)

- Corpus Escrito

- Textos Jornalísticos
- Textos de Blogs
- Textos Científicos

- Palavras foram transcritas, etiquetadas e acentuadas automaticamente

# Dados do Corpus ABG

## Corpus Escrito

<b>Estadão</b>	397.869
<b>Artigos</b>	342.871
<b>Blogs</b>	215.126

<b>Folha</b>	819.381
--------------	---------

## TOTAIS PARCIAIS

Ocorrências	1.775.247
Tipos	104.364

## TOTAL GERAL

<b>Ocorrências</b>	
<b>Tipos</b>	

## Corpus Oral

<b>C-oral</b>	12.079
<b>Iboruna</b>	734.991
<b>ProjetoSP2010</b>	230.116

<b>ProjetoSP+</b>	986.612
-------------------	---------

Ocorrências	1.963.798
Tipos	40.586

3.739.045
123.245

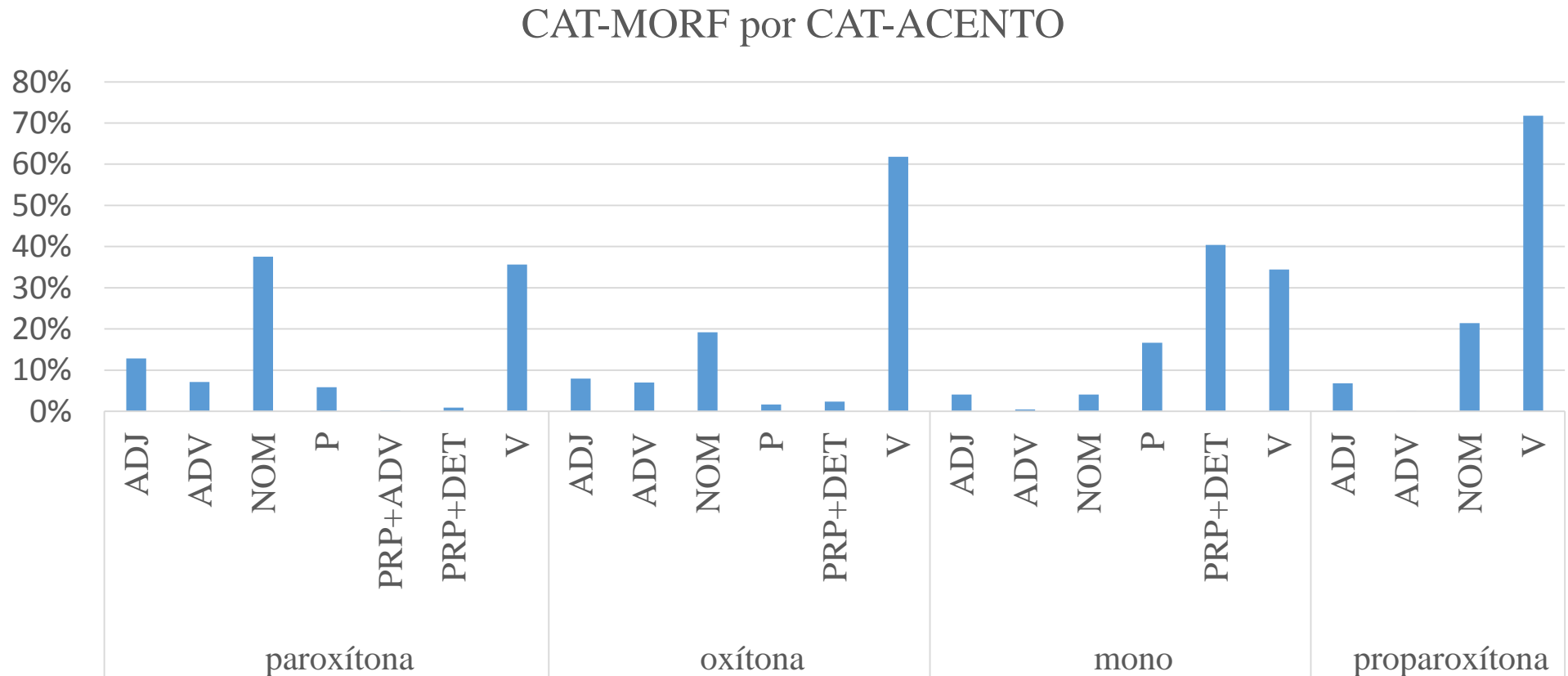
# Corpus ABG- variáveis

Varíaveis	Tipo	Variantes
Categoria Acentual (D)	Nominal	Oxítona Paroxítona Proparoxítona
Classe Morfossintática	Nominal	NOM, V, ADJ...
Segmento Final	Nominal	R,S,V,N
Origem Etimológica	Nominal	Latina, Grega, Indígena, Africana, Inglesa
Nº Ocorrências no Corpus	Numérica	-
Tamanho da palavra	Numérica	Número de sílabas



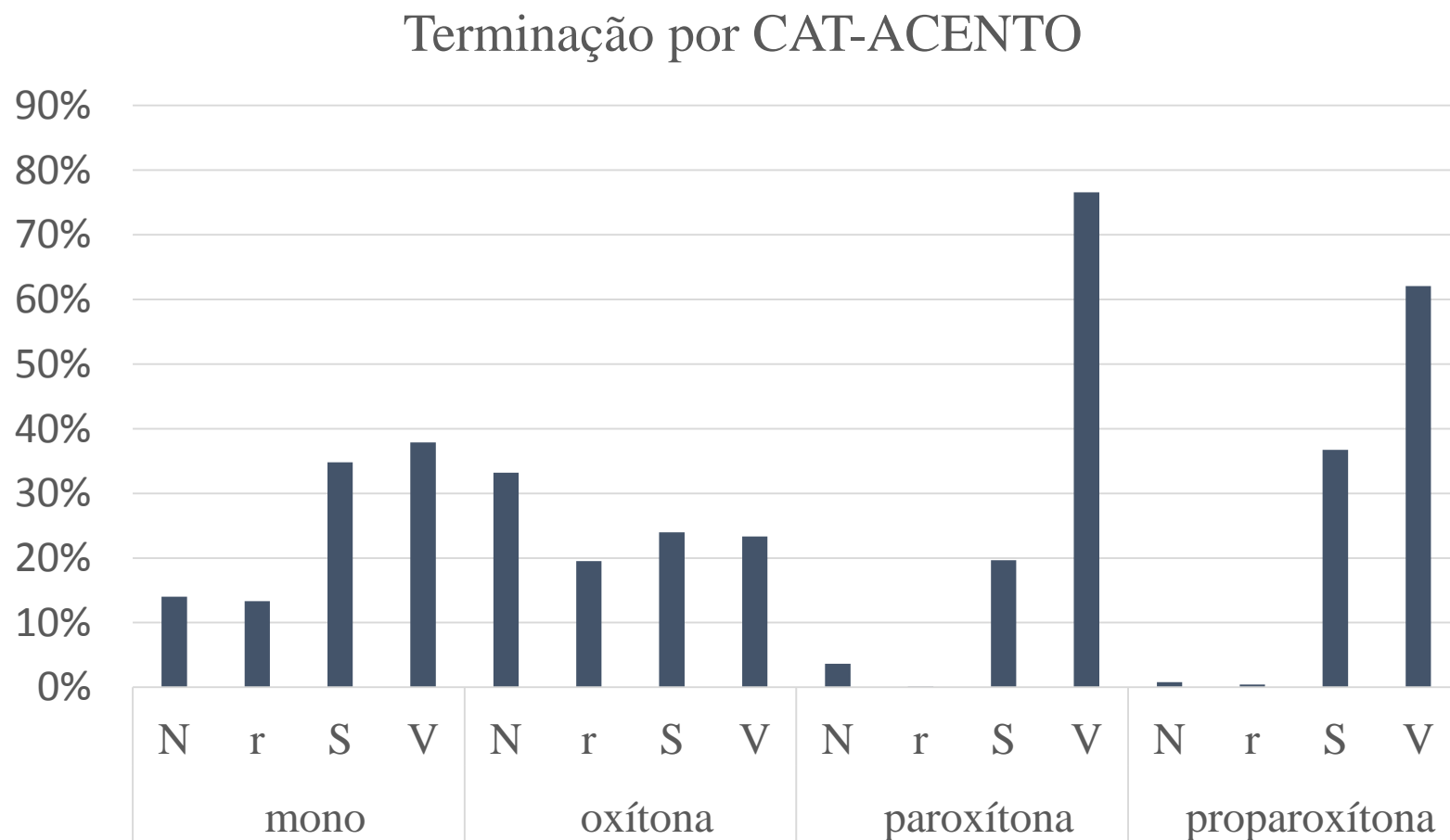
# Acento – Análise descritiva

- Qual a relação entre categorias morfológicas e categoria acentual?



# Acento – Análise descritiva: peso

- Existe relação entre o segmento final da palavra e a categoria acentual?



# Conclusão – Encaminhamentos

- H1: Há variáveis (linguísticas e não são arbitrárias) além da categoria morfosintática e do peso silábico que estão relacionadas ao acento.
- H1a: A frequência de uma sílaba está relacionada com o acento.  
(Frequência da sílaba como tônica e átona?)
- H1b: A origem do radical da palavra é uma variável relacionada ao acento.