

RELATÓRIO DE ANÁLISE ESTATÍSTICA SOBRE O PROJETO:

“O processamento de pseudopalavras no Português Brasileiro”

Giovanna Vendeiro Vilar

Mariana Almeida

Renata Massami Hirota

Viviana Giampaoli

São Paulo, maio de 2021

CENTRO DE ESTATÍSTICA APLICADA - CEA - USP

TÍTULO: Relatório de análise estatística sobre o projeto: “O processamento de pseudopalavras no Português Brasileiro”

PESQUISADORA: Aline Benevides

ORIENTADORA: Profa. Dra. Raquel Santana Santos

INSTITUIÇÃO: Faculdade de Filosofia e Ciências Humanas da Universidade de São Paulo

FINALIDADE DO PROJETO: Doutorado

RESPONSÁVEIS PELA ANÁLISE:

Giovanna Vendeiro Vilar

Mariana Almeida

Renata Massami Hirota

Viviana Giampaoli

REFERÊNCIA DESTE TRABALHO: ALMEIDA, M.C.; GIAMPAOLI, V.; HIROTA, R.M.; VILAR, G.V. **Relatório de análise estatística sobre o projeto: “O processamento de pseudopalavras no Português Brasileiro”**. São Paulo, IME-USP, ano. (RAE–CEA-21P02)

Ficha técnica

Referências bibliográficas

COLLISCHONN, G. 1999. *Acento Em Português. In BISOL, I. (ed.) Introdução a Estudos de Fonologia Do Português Brasileiro*. 1.ed ed. Porto Alegre: EDIPUCRS.

FIORIN, J. L. 2019. *Linguística? Que é Isso?* 1.ed ed. São Paulo: Contexto.

PETTER, M. 2007. *Introdução à Linguística i: Objetos Teóricos*. 5.ed ed. São Paulo: Contexto.

PRIBERAN. n.d. ""pseudopalavra", in Dicionário Priberam Da língua Portuguesa." <https://dicionario.priberam.org/pseudopalavra>.

Programas computacionais utilizados

R 3.0.4

Técnicas estatísticas utilizadas

Área de aplicação

Linguística (14:110)

Sumário

1 Introdução	1
2 Objetivo	2
3 Descrição do estudo	2
3.1 Limitações do estudo	3
3.2 Conceitos do estudo e Validação	4
4 Descrição das variáveis	5
4.1 Variável Dependente	5
4.2 Variáveis Linguísticas	5
4.3 Variáveis Extralinguísticas	6
4.4 Variáveis Experimentais	6
5 Análise descritiva	7
5.1 Perfil dos participantes	7
5.1.1 Sexo e idade	7
5.1.2 Naturalidade	8
5.1.3 Área de formação	8
5.2 Variáveis linguísticas	8
5.3 Conclusão da análise descritiva e próximos passos	10
Apêndices	
A Tabelas	12

1 Introdução

De acordo com PETTER (2007), o interesse pela linguagem é antigo e vem sendo expresso por meio de mitos, lendas, cantos, rituais e trabalhos eruditos que buscam conhecer essa capacidade humana como sistema de comunicação. A autora aponta que, a partir do século XX, os estudos linguísticos passaram a ter um caráter científico, ou seja, centrados na observação dos fatos a partir de pressupostos teóricos da linguagem, no estabelecimento de hipóteses e na examinação mediante experimentos.

Conforme descrito por FIORIN (2019), a linguística é uma ciência da linguagem porque, ao contrário da gramática, ela tem como objetivo estabelecer o que uma língua é e por que é de uma determinada maneira. Logo, a área estuda os aspectos fonéticos, morfológicos, sintáticos, semânticos, sociais e psicológicos de uma língua, e neste caso, o português brasileiro. Dentro deste contexto, existe o conceito de *pseudopalavra*, que, de acordo com o dicionário PRIBERAN (n.d.), é uma

Sequência regular e pronunciável de caracteres que não tem um significado numa língua, apesar de obedecer às regras ortográficas, morfológicas ou de pronúncia

No português brasileiro existem três classes de palavras segundo sua tonicidade: oxítona, paroxítona e proparoxítona. Essas denominações estão relacionadas à intensidade dada a determinadas sílabas na pronúncia das palavras. Aquela que é pronunciada de forma mais acentuada é a sílaba tônica. Assim, oxítonas são as palavras cuja sílaba tônica é a última; paroxítonas são as palavras cuja sílaba tônica é a penúltima; e proparoxítonas são as palavras cuja sílaba tônica é a antepenúltima.

O intuito do trabalho é investigar a maneira como os falantes nativos do português atribuem a tonicidade em pseudopalavras parcialmente relacionadas a vocábulos existentes no idioma. Em outras palavras, busca-se compreender como o indivíduo, ao se deparar com uma palavra nova, nesse caso, uma pseudopalavra, define a sílaba tônica. Além disso, busca-se entender quais são os outros fatores, tais como

os conhecimentos linguísticos do falante e as estruturas linguísticas das palavras que podem influenciar nesse processo de classificação e portanto da determinação da entonação.

2 Objetivo

O objetivo principal do trabalho é verificar se pseudopalavras que se assemelham a palavras reais podem sofrer um processo análogo e ter o mesmo padrão acentual da referência real. Além disso, buscamos entender quais são os conhecimentos linguísticos do falante utilizados nesse processo de acentuação tônica das pseudopalavras e determinar um modelo que explique a classificação das pseudopalavras.

Algumas perguntas a serem respondidas pela análise estatística são:

1. Os grupos experimentais têm influencia na acentuação tônica das pseudopalavras?
2. A taxa de similaridade influencia na associação acentual? Entender se quanto mais similar a pseudopalavra for da palavra real, maiores são as chances de atribuição do mesmo padrão acentual da palavra alvo
3. As variáveis selecionadas pelo modelo estão em concordância com a literatura da área?
4. Há associação entre graduação em letras e a classificação tônica da pseudopalavra? Há associação entre conhecimento em música e a classificação tônica da pseudopalavra

3 Descrição do estudo

O estudo foi realizado de maneira remota com reunião individuais da pesquisadora e cada um dos participantes via Google Meet. Consistiu em apresentar aos participantes, através do software Psychopy, 372 pseudopalavras agrupadas nos denominados grupos de estímulos e registrar a forma como eles reproduziam

paroxítona e proparoxítona. Essa é a variável dependente do estudo, que será descrita junto às outras variáveis na seção 4.

A coleta dos dados foi realizada no início do primeiro semestre de 2020 com 34 indivíduos que, por meio de divulgações em redes sociais e de colegas, se voluntariaram a participar do experimento. Os voluntários tiveram como pré-requisitos, ser maior de 18 anos, ser falante nativo do português brasileiro e não ter estudado linguística.

Entre os participantes da pesquisa encontram-se estudantes do primeiro semestre da faculdade de Letras da Universidade de São Paulo, músicos, alguns residentes de fora do estado de São Paulo, entre outros. Supõe-se que os alunos do primeiro semestre do curso de Letras ainda não têm conhecimento na área. O perfil dos participantes é descrito na seção 5.

3.1 Limitações do estudo

Identificamos dois eventuais problemas na coleta de dados que tentaremos contornar nas análises e que são descritos a seguir.

Problemas técnicos e interferência externa:

Destaca-se a perda de algumas respostas durante o processo de coleta de dados, visto que ruídos externos impediram que algumas entonações fossem captadas e registradas na gravação. Logo, na base de dados não temos 372 registros de pseudopalavras para todos os participantes. Portanto, um total de 12.511 dados serão utilizados na análise, em vez dos 12.648 esperados, o qual não representa uma perda substancial.

Problemas de aleatorização:

As palavras foram aleatorizadas no Excel e divididas em 4 conjuntos, a serem apresentados aos participantes com um intervalo de tempo entre cada conjunto. Por limitações do software, a ordem de apresentação desses conjuntos não pôde ser aleatorizada, apenas a ordem das palavras dentro de cada conjunto.

No entanto, para alguns participantes o software Psychopy apresentou problemas e eles tiveram que continuar o experimento a partir de slides com uma ordem aleatória pré-estabelecida. Em outras palavras, todos os indivíduos que em algum momento acompanharam o experimento pelos slides seguiram com palavras apresentadas na mesma ordem (a primeira aleatorização retirada do Excel).

Ressaltamos que essa variável não foi controlada durante a coleta de dados, ou seja, não foi definido uma quantidade de palavras de cada categoria em cada bloco.

3.2 Conceitos do estudo e Validação

O conceito de similaridade foi construído a partir de uma taxa de similaridade definida com base nas mudanças feitas na palavra-alvo até a obtenção da pseudopalavra. De acordo com o tipo e a quantidade de alterações, foi estabelecido um valor de 1 a 10 -taxa de similaridade-, onde valores menores de 4 determinam palavras similares e valores maiores ou iguais a 5 determinam palavras dissimilares. Logo, foi necessário definir um modo de entender se essa classificação em similar e dissimilar é válida.

Esse ponto está relacionado aos testes para validar se a pseudopalavra é similar a palavra a partir da qual ela foi criada (palavra alvo). Nessa etapa -chamada de validação- pediu-se para 10 falantes do português, que não fazem parte do estudo final, listarem a palavra real a qual eles associavam a palavra criada. Considerou-se como validadas as pseudopalavras cuja associação foi a palavra alvo na resposta de, no mínimo, oito indivíduos. Porém, pseudopalavras nas quais sete falantes apresentaram a associação correta foram classificadas como quase validadas.

Da mesma forma, precisou-se elaborar um modo de validar se a pseudopalavra realmente é dissimilar à palavra alvo. Foram consideradas validadas como dissimilares à palavra alvo pseudopalavras não associadas a uma mesma resposta por mais de dois indivíduos. Em outras palavras, se mais de dois falantes lembraram de uma mesma palavra ao ler a pseudopalavra, ela foi considerada dissimilar à sua palavra alvo - pois entendeu-se que ela está “parecida” com uma palavra real. Pseudopalavras nas quais três falantes apresentaram a mesma associação foram

classificadas como quase validadas. Ressalta-se que, nesse caso, a associação não necessariamente precisava ser com a palavra alvo.

4 Descrição das variáveis

Foram coletadas variáveis linguísticas -relacionadas às pseudopalavras-, extralinguísticas -relacionadas aos participantes- e experimentais -relacionadas ao estudo-, que podem, segundo a literatura da área, influenciar o comportamento acentual no português. A seguir, listamos as variáveis pré-selecionadas para o estudo.

4.1 Variável Dependente

A variável resposta de interesse é **Tonicidade da pseudopalavra**, ou seja, a classificação acentual tônica da pseudopalavra (oxítona, paroxítona e proparoxítona)

4.2 Variáveis Linguísticas

As palavras que deram origem às pseudopalavras, definidas como palavras alvo, foram classificadas em dois níveis de acordo com a sua ocorrência no Corpus brasileiro, corpus linguístico coordenado pelo pesquisador Antonio Paulo Berber Sardinha. Se a palavra possui mais de 100 mil ocorrências no corpus ela é classificada como alta frequência e se possui menos de 2 mil ocorrências ela é classificada como baixa frequência. A junção da ideia de frequência e similaridade entre a palavra alvo e a pseudopalavra resultou na criação de uma variável com 4 categorias chamada grupo de estímulos.

Ademais, as pseudopalavras foram construídas com três sílabas de extensão para que os três padrões acentuais do português brasileiro pudessem ser produzidos.

- **Validação:** s = sim, n = não validada e q = quase validada
- **Taxa de validação:** indica quantas pessoas do estudo preliminar informaram que a palavra era similar ou dissimilar
- **Palavra alvo:** palavra real que deu origem à pseudopalavra

Tonicidade da palavra alvo: oxítona, paroxítona e proparoxítona

- **Estrutura da palavra alvo:** indica qual é a estrutura da pseudopalavra (CV-CV-CV ou CV-CV-CVC)
- **Pseudopalavra:** refere-se a cada um dos estímulos criados
- **Segmento modificado:** indica qual letra foi modificada na criação da pseudopalavra a partir da palavra real (consoante ou vogal)
- **Grupo dos estímulos:** indica o efeito da similaridade (entre a pseudopalavra e a palavra real) e da frequência (alta e baixa) na produção acentual
 - 1 = pseudopalavras similares de alta frequência
 - 2 = pseudopalavras dissimilares de alta frequência
 - 3 = pseudopalavras similares de baixa frequência
 - 4 = pseudopalavras dissimilares de baixa frequência
- **Taxa de similaridade:** 1, 2, 3 (grupos similares), 5, 6, 7, 8, 9, 10 (grupos dissimilares)

4.3 Variáveis Extralinguísticas

- **Participante:** indica os 34 participantes do experimento
- **Idade:** de 18 a 60 (anos)
- **Gênero:** feminino e masculino
- **Naturalidade:** indica a cidade em que o participante nasceu
- **Escolaridade:** ensino fundamental a mestrado
- **Área de formação:** 0 = outros e 1 = letras
- **Línguas:** 0 = não tem conhecimento em línguas e 1 = tem conhecimento em línguas
- **Música:** 0 = não tem conhecimento em música e 1 = tem conhecimento em música

4.4 Variáveis Experimentais

Inicialmente, as pseudopalavras foram aleatorizadas de forma geral no software Excel (ALEATORIOENTRE) e separadas em 4 blocos com 93 pseudopalavras cada, resultando em um total de 372 palavras criadas.

- **Bloco de apresentação:** indica em qual bloco (ou rotina) a pseudopalavra foi inserida (1, 2, 3 ou 4)
- **Ordem de apresentação:** indica em qual ordem a pseudopalavra foi apresentada dentro do bloco de apresentação (1 a 93). Para os indivíduos que fizeram toda a dinâmica no software, a ordem das pseudopalavras era diferente dentro de cada bloco.
- **Aleatorização:** codifica se o bloco de apresentação foi aleatorizado para o indivíduo ou se foi a aleatorização prévia (s = o estímulo foi aleatorizado e n = não houve aleatorização)

Inicialmente, as pseudopalavras foram aleatorizadas de forma geral no software Excel (ALEATORIOENTRE) e separadas em 4 blocos com 93 pseudopalavras cada, resultando em um total de 372 palavras criadas. Dentro de cada bloco, a ordem em que as pseudopalavras foram apresentadas aos falantes também foi aleatorizada, porém, para alguns participantes o software Psychopy apresentou problemas e eles tiveram que continuar o experimento a partir de slides com uma ordem aleatória pré-estabelecida.

14 participantes em algum momento continuaram o experimento via slide

Em outras palavras, todos os indivíduos que em algum momento acompanharam o experimento pelos slides seguiram com palavras apresentadas na mesma ordem (a primeira aleatorização retirada do Excel).

5 Análise descritiva

5.1 Perfil dos participantes

5.1.1 Sexo e idade

Os 34 participantes do estudo estão divididos entre 21 mulheres e 13 homens, com idades que variam entre 18 e 60 anos. A Figura A.1 mostra a distribuição dos respondentes segundo a faixa etária e gênero. A faixa etária foi dividida a partir dos quartis da variável idade, para resumir as informações da amostra.

5.1.2 Naturalidade

A Tabela [A.1](#) exibe o perfil dos informantes segundo sua naturalidade (variável agrupada em razão da baixa volumetria de indivíduos por UF). É possível perceber que a maior proporção (82%) dos participantes reside no estado de São Paulo, portanto, essa variável não será considerada na análise.

5.1.3 Área de formação

Parte considerável dos participantes (10) são ingressantes do curso de Letras, logo, essa variável foi categorizada em dois níveis. A Tabela [A.2](#) mostra esse agrupamento da formação dos voluntários. A influência do curso de Letras na categorização das pseudopalavras é um dos fatores de possível interesse na análise dos dados.

Outras variáveis relacionadas com a linguagem (línguas e música) também foram transformadas em variáveis binárias, pois há interesse em entender se o conhecimento nessas áreas afeta na percepção da tonicidade de palavras. A distribuição das variáveis já agrupadas também pode ser observada nas Tabelas [A.3](#) e [A.4](#), respectivamente.

5.2 Variáveis linguísticas

A tonacidade de produção do tipo proparoxítona é a menos expressiva na base, enquanto a categoria paroxítona (penúltima sílaba é tônica) aparece com maior frequência nas respostas dos participantes, como visto na Tabela [5.1](#) a seguir.

Portanto, o processo de acentuação das pseudopalavras nesse estudo parece ir de acordo com a afirmação de [COLLISCHONN \(1999\)](#) de que

Podemos considerar que o acento proparoxítono é marcado, no sentido de que é menos usual. É um acento especial, contrário à tendência geral de acentuar a penúltima sílaba

Tabela 5.1: Distribuição das respostas nos níveis da variável Tonicidade de produção.

Tonicidade de produção	Total
oxítona	4904
paroxítona	7163
proparoxítona	444

Além disso, na Tabela [A.5](#) verifica-se que a proporção de palavras alvo proparoxítonas é inferior às outras categorias em todos os grupos, o que já era esperado devido à baixa frequência de palavras proparoxítonas de três sílabas na estrutura proposta (CV-CV-CV ou CV-CV-CVC, C = consoante e V = vogal).

Observa-se na [A.6](#) que os grupos 2 e 4 (ou seja, os grupos em que as palavras sofreram mais alterações em relação à referência original) apresentam uma proporção menor de proparoxítonas produzidas quando comparados aos grupos 1 e 3.

Comparando a tonicidade das pseudopalavras e a tonicidade das palavras-alvo (vide Tabela [5.2](#)), nota-se que 73% das pseudopalavras classificadas como proparoxítonas tinham uma palavra-alvo de tonicidade proparoxítona.

Entretanto, cerca de 77% das pseudopalavras que têm como alvo uma palavra proparoxítona foram categorizadas pelos participantes como paroxítonas (Tabela [A.7](#)).

Logo, pode-se pensar que pseudopalavras derivadas de proparoxítonas não tendem a ser acentuadas com a mesma tonicidade da palavra alvo, mas pseudopalavras lidas como proparoxítonas tender a ser derivadas de proparoxítonas

Tabela 5.2: Frequência da tonicidade das pseudopalavras por nível de tonicidade das palavras-alvo.

Tonicidade produção	Tonicidade da palavra-alvo			Total
	oxítona	paroxítona	proparoxítona	
oxítona	2642 (54%)	1939 (39.54%)	323 (7%)	4904 (100%)
paroxítona	2154 (30%)	2824 (39.42%)	2185 (31%)	7163 (100%)
proparoxítona	49 (11%)	72 (16.22%)	323 (73%)	444 (100%)

Na Tabela 5.3 vemos que aproximadamente 73% das pseudopalavras foram validadas de acordo com o processo descrito anteriormente no capítulo 3 (Descrição). Porém, cerca de 81% das não validadas foram classificadas como paroxítonas, e mais de 55% pertencem ao grupo das similares (A.8).

Já na Tabela A.8 percebe-se que nos grupos 1 (pseudopalavras similares de alta frequência) e 3 (pseudopalavras similares de baixa frequência) a porcentagem de palavras não validadas é superior aos demais grupos, o que parece indicar que a validação de palavras similares é mais complexa e difícil.

Tabela 5.3: Frequência da tonicidade das pseudopalavras, por nível da variável Validação.

Tonicidade produção	Status Validação			Total
	Não Validada	Quase Validada	Validada	
oxítona	335 (6.8%)	529 (10.787%)	4040 (82.4%)	4904 (100%)
paroxítona	1727 (24.1%)	771 (10.764%)	4665 (65.1%)	7163 (100%)
proparoxítona	56 (12.6%)	48 (10.811%)	340 (76.6%)	444 (100%)

Por fim, podemos ver que a divisão de estrutura da palavra-alvo é bem distribuída, com exceção da estrutura CV-CV-CVC em proparoxítonas, o que já era esperado, uma vez que a frequência de palavras com essas características no *Corpus Brasileiro* é baixa (vide Tabela 5.4).

Tabela 5.4: Frequência da Tonicidade alvo por Estrutura das palavras.

Tonicidade palavra-alvo	Estrutura da palavra		Total
	CV-CV-CV	CV-CV-CVC	
oxítona	2417 (49.89%)	2428 (50.11%)	4845
paroxítona	2427 (50.20%)	2408 (49.80%)	4835
proparoxítona	2427 (85.73%)	404 (14.27%)	2831

5.3 Conclusão da análise descritiva e próximos passos

A partir dos resultados presentes nas tabelas da seção anterior, percebe-se um comportamento não homogêneo entre níveis de algumas variáveis. Portanto, o próximo passo do estudo consiste em realizar uma análise univariada, com testes qui-quadrado, que permite analisar a relação de independência entre variáveis qualitativas. Algumas hipóteses iniciais são:

H_0 : O grupo experimental determina a localização acentual

H_1 : O grupo experimental não determina a localização acentual

2. Observar se a tonicidade de produção é coincidente com a tonicidade alvo

H_0 : Há associação entre os padrões acentuais

H_1 : Não há associação entre os padrões acentuais

3. Entender se a tonicidade da palavra real influencia mais em pseudopalavras mais similares e menos em pseudopalavras mais dissimilares

H_0 : A taxa de similaridade não influencia na associação acentual

H_1 : A taxa de similaridade influencia na associação acentual

4. A validação da similaridade e dissimilaridade em relação à palavra alvo afeta a acentuação tônica das pseudopalavras?

H_0 : Pseudopalavras não validadas têm o mesmo comportamento das validadas e quase validadas

H_1 : Pseudopalavras não validadas têm comportamento distinto das validadas e quase validadas

5. Entender se o conhecimento em Música influencia na acentuação tônica

H_0 : Não há associação entre conhecimento em Música e a tonicidade produzida

H_1 : Há associação entre conhecimento em Música e a tonicidade produzida

6. Entender se a estudantes no início da graduação em Letras tem um comportamento diferente no processo de acentuação tônica

H_0 : Não há associação entre graduação em Letras e a tonicidade produzida

H_1 : Há associação entre graduação em Letras e a tonicidade produzida

Além disso, a parte inferencial da análise será pautada em um modelo hierárquico no qual os participantes representam um nível com efeito aleatório, e a variável dependente será a tonicidade de produção em cada pseudopalavra.

Apêndices

A Tabelas

Tabela A.1: Perfil dos participantes de acordo com a Naturalidade.

Naturalidade	Total
São Paulo, SP	16
Outros municípios de SP	12
Outras UF	6

Tabela A.2: Frequência da variável dicotômica Área de formação.

Área de formação	Total
Outro	19
Letras	15

Tabela A.3: Frequência da variável dicotômica Línguas.

Conhecimento em línguas	Total
1	26
0	8

Tabela A.4: Frequência da variável dicotômica Música.

Conhecimento em música	Total
1	21
0	13

Tabela A.5: Frequência do Grupo experimental por Tonicidade da palavra-alvo.

Grupo	Tonicidade da palavra alvo			Total
	oxítona	paroxítona	proparoxítona	
1 (Similar de alta freq.)	1213 (40.0%)	1216 (40.1%)	606 (20.0%)	3035 (100.0%)
2 (Dissimilar de alta freq.)	1207 (40.0%)	1205 (39.9%)	606 (20.1%)	3018 (100.0%)
3 (Similar de baixa freq.)	1216 (37.6%)	1209 (37.4%)	810 (25.0%)	3235 (100.0%)
4 (Dissimilar de baixa freq.)	1209 (37.5%)	1205 (37.4%)	809 (25.1%)	3223 (100.0%)

Tabela A.6: Frequência do Grupo experimental por Tonicidade de produção (da pseudopalavra).

Tonicidade produção	Grupo				Total
	1 (Similar de alta freq.)	2 (Dissimilar de alta freq.)	3 (Similar de baixa freq.)	4 (Dissimilar de baixa freq.)	
oxítona	1077 (22.0%)	1152 (23.5%)	1308 (26.7%)	1367 (27.9%)	4904
paroxítona	1805 (25.2%)	1818 (25.4%)	1760 (24.6%)	1780 (24.8%)	7163
proparoxítona	153 (34.5%)	48 (10.8%)	167 (37.6%)	76 (17.1%)	444

Tabela A.7: Frequência da tonicidade das palavras-alvo por nível de tonicidade das pseudopalavras.

Tonicidade da pseudopalavra				
tonicidade_alvo	oxitona	paroxitona	proparoxitona	total
oxítona	2642 (55%)	2154 (44%)	49 (1.01%)	4845 (100%)
paroxítona	1939 (40%)	2824 (58%)	72 (1.49%)	4835 (100%)
proparoxítona	323 (11%)	2185 (77%)	323 (11.41%)	2831 (100%)

Tabela A.8: Frequência das pseudopalavras nos Grupos, por status de Validação.

Status Validação				
Grupo	Não Validada	Quase Validada	Validada	Total
1	772 (25.44%)	202 (6.66%)	2061 (67.9%)	3035 (100%)
2	135 (4.47%)	372 (12.33%)	2511 (83.2%)	3018 (100%)
3	1044 (32.27%)	203 (6.28%)	1988 (61.5%)	3235 (100%)
4	167 (5.18%)	571 (17.72%)	2485 (77.1%)	3223 (100%)

Tabela A.9: Atribuições acentuais por Estrutura das palavras.

Estrutura da palavra			
Tonicidade palavra produção	CV-CV-CV	CV-CV-CVC	Total
oxítona	748 (15%)	4156 (85%)	4904
paroxítona	6200 (87%)	963 (13%)	7163
proparoxítona	323 (73%)	121 (27%)	444