

QUANTIFICAÇÃO DAS INTERPRETAÇÕES DE 'SOZINHO' NO PB: UMA ANÁLISE COMPARATIVA DE TEXTOS ANTERIORES E POSTERIORES À 1900

ELISA RODRIGUES

RESUMO: O presente trabalho trata-se de um projeto desenvolvido na disciplina ‘Laboratório 7: Processamento de Linguagem Natural’, ofertado no período 2024/1 pelo Departamento de Letras para o curso de Bacharelado em Linguística e sob orientação do Prof. Dr. Pablo Arantes. Neste projeto, foi desenvolvido um script em Linguagem R que processa textos de dois períodos diferentes como dados de entrada buscando por sentenças que contenham a palavra 'sozinho' e suas variações em número e gênero, bem como suas variações ortográficas. A partir da coleta das sentenças, cada uma delas foi rotulada manualmente como uma das 7 interpretações possíveis ativadas por 'sozinho'. Por fim, foi gerado um gráfico de barras pela Linguagem R a partir de um arquivo com as sentenças rotuladas para os textos de cada período. Desse modo, foi possível realizar uma análise quantitativa sobre as ocorrências das interpretações de 'sozinho' para os textos de cada período, assim testando a hipótese principal: interpretações eventivas de 'sozinho' ocorrem em apenas textos de um período.

INTRODUÇÃO

O item ‘sozinho’ do português brasileiro (PB), bem como suas contrapartes em outras línguas naturais, é pouco estudado no âmbito das teorias formais, ainda mais considerando o quadro da Semântica Formal. Exceções seriam trabalhos como Basso & Palmieri (2021) e Basso & Rodrigues (2024), que fornecem uma descrição de diferentes interpretações e propõem uma análise (anti-)comitativa¹ para uma explicação semântica unificada de ‘sozinho’. Em específico, Basso & Rodrigues (2024) descrevem sete interpretações distintas ativadas pelo item ‘sozinho’: espaço-temporal, modificador argumental, (anti-)causal,

¹ Comitatividade é uma categoria gramatical que expressa semanticamente a noção de companhia, em que, para diferentes níveis de interpretação, resulta em usos de instrumentalidade, modalidade e acompanhamento. Em PB, a preposição ‘com’ é um exemplo característico de item comitativo:

a. Cortei a maçã com a faca. – *instrumental*

b. Maria gritou com raiva. – *modal*

c. Maria viajou com João. – *acompanhamento*

De acordo com Basso & Rodrigues (2024), o item ‘sozinho’ seria um modificador de (anti-)comitatividade, expressando semanticamente a noção de não possuir companhia.

autônoma, emocional, comportamental e mereológica. Os autores identificam que é possível categorizar todas as interpretações de ‘sozinho’ em três categorias a partir do seu escopo na sentença, conforme mostra a tabela a seguir.

indivíduo como escopo	papel temático como escopo	evento como escopo
Maria tá sozinha na França <i>sozinho espaço-temporal</i>	Maria comeu o bolo sozinho <i>sozinho modificador argumental</i> (argumento interno)	A TV desligou sozinha <i>sozinho (anti-)causal</i>
Achei a/uma meia sozinha <i>sozinho mereológico</i>	Maria comeu o bolo sozinha <i>sozinho modificador argumental</i> (argumento externo)	A garrafa para em pé sozinha <i>sozinho (anti-)assistido</i>
Ana tá se sentindo sozinha <i>sozinho emocional</i>		
João é sozinho <i>sozinho comportamental</i>		

Tabela 1 – As três categorias para as interpretações de 'sozinho' a partir de seu escopo na sentença.

Para o presente trabalho, lido com as interpretações eventivas de ‘sozinho’, que são as interpretações (anti-)causal e autônoma. Até então, não foi observado na literatura especializada nenhuma menção de ‘sozinho’ ou suas contrapartes em outras línguas naturais que atuem sobre o evento da sentença. Minha hipótese é a de que a modificação de eventos seja um comportamento linguístico recentemente caracterizado pelo item ‘sozinho’, portanto, não deve ser possível encontrar as interpretações (anti-)causal e autônoma de ‘sozinho’ em textos escritos em um PB atual. Desse modo, será realizada uma análise estatística e comparativa sobre a ocorrência de interpretações de ‘sozinho’ em textos escritos em PB atual e não-atual. O trabalho, então, está dividido nos tópicos a seguir: (i) sobre o fenômeno linguístico; (ii) sobre os materiais utilizados para a análise; (iii) sobre a investigação para alcançar os textos; (iv) sobre o script desenvolvido; (v) análise estatística e resultados.

1. SOBRE O FENÔMENO LINGUÍSTICO: 'SOZINHO' EVENTIVO

Conforme apresentado anteriormente, o fenômeno do 'sozinho' eventivo se trata das interpretações de 'sozinho' que possuem como escopo da sentença entidades do tipo evento, compreendendo as interpretações (anti-)causal e (anti-)assistida. Neste tópico, então, apresento o comportamento linguístico do 'sozinho' eventivo levando em conta suas características sintáticas e semânticas.

1.1 'SOZINHO' COMO MODIFICADOR DE VP

A palavra 'sozinho' opera sobre diferentes tipos de construções sintáticas. Para o caso do 'sozinho' eventivo, 'sozinho' atua como um modificador de Sintagma Verbal (VP). A modificação de VP pelo item 'sozinho' ocorre através de construções adverbiais adjuntivas, em que a ausência de 'sozinho' não afeta a gramaticalidade da sentença:

- (1) a. A porta bateu sozinha.
- b. O bebê anda sozinho.
- c. Pedro apresentou sozinho o trabalho.
- d. Os funcionários carregam sozinhos as caixas.

Observamos que nas construções adverbiais há concordância em gênero e número do 'sozinho' com o argumento externo do predicado verbal modificado por ele. Além disso, as construções adverbiais de 'sozinho' ocorrem tanto em posição pós-verbal quanto pré-verbal:

- (2) a. A TV desligou sozinha.
- b. A TV sozinha desligou.
- c. Os funcionários carregam sozinhos as caixas.
- d. Os funcionários sozinhos carregam as caixas.

No entanto, a posição pré-verbal das construções adverbiais do 'sozinho' modificador de VP ocorre apenas na presença de uma curva prosódica específica sobre o item 'sozinho'. Notamos também que as construções adverbiais de 'sozinho' não permitem a combinação com intensificadores:

- (3) a. *A TV desligou muito sozinha.
b. #Os funcionários carregam muito sozinhos as caixas.

Estruturas comparativas também geram sentenças agramaticais ou pragmaticamente estranhas quando combinadas com as construções adverbiais de 'sozinho':

- (4) a. *A porta bateu mais sozinha que a janela.
b. #Pedro apresentou mais sozinho o trabalho que Joana.

Algumas sentenças geradas com o item 'sozinho' são sentenças ambíguas, em que a disputa da ambiguidade está em processar a estrutura de 'sozinho' como uma construção adverbial ou como uma construção predicativa secundária, como ilustrado a seguir:

- (5) a. Pedro apresentou sozinho o trabalho. (*adverbial* ou *predicativa secundária*)
b. Pedro apresentou muito sozinho o trabalho. (*predicativa secundária*)

Em (5a), há uma ambiguidade sintática em que 'sozinho' pode ser processado tanto como um modificador de VP em 'apresentou' quanto como um modificador de NP em 'Pedro'. Como vimos, construções adverbiais de 'sozinho' não permitem a combinação com intensificadores. Desse modo, em (5b), quando 'sozinho' se combina com 'muito', a ambiguidade sintática é desfeita e 'sozinho' passa a ser processado como um modificador de NP – assim tornando a sentença (5b) gramatical, uma vez que intensificadores podem se combinar com o 'sozinho' modificador de NP.

1.2 EVENTO COMO ESCOPO

Nas construções em que 'sozinho' atua como um modificador de VP, o evento do verbo núcleo do sintagma modificado é tomado como escopo de 'sozinho', resultando nas interpretações eventivas de (anti-)causalidade e (anti-)assistência. Semanticamente, as interpretações (anti-)causal e (anti-)assistidas de 'sozinho' se diferenciam a partir da natureza do evento e dos participantes que compõem ele.

- (6) a. A TV desligou sozinha. – (*anti-*)*causal*
b. O bebê anda em pé sozinho. – (*anti-*)*assistido*

Em (6a), a interpretação ativada é a de que não há causa aparente ou identificável para o falante explicar porque a TV desligou – ela desligou "do nada". Em (6b), a interpretação ativada é a de que o bebê possui a capacidade de realizar o evento de andar em pé por conta própria – é mérito do bebê conseguir andar em pé. Para diferenciar as interpretações eventivas de 'sozinho', é possível realizar testes linguísticos com analíticos causativos:

- (7) a. *A TV desligou sozinha quando eu tirei da tomada.
b. O bebê anda sozinho quando eu dou um empurrãozinho nas costas dele.

As interpretações (anti-)causais de 'sozinho' não permitem a combinação com analíticos causativos, como demonstrado em (7a). Dado que *e* corresponde ao evento da TV desligar, não pode haver nenhum evento *e'* tal que $e' < e$ e *e'* apresente uma causa para *e*. Ou seja, não deve haver nenhum pré-evento que seja uma explicação possível para o evento modificado por 'sozinho' ocorrer. Por outro lado, para as interpretações (anti-)assistidas de 'sozinho', é possível identificar um pré-evento que justifique o evento modificado por 'sozinho' na sentença, como demonstrado em (7b). Dado que *e* corresponde ao evento de andar em pé, mesmo que haja um evento *e'* que antecede *e*, apresentando uma possível causa para *e*, continua sendo mérito do participante-agente (bebê) de *e* a capacidade de realizar o evento.

Outras línguas naturais também apresentam interpretações (anti-)causais e (anti-)assistidas, como as ativadas pelas palavras *afto* do grego e *auto* do francês. No entanto, o PB veicula essas interpretações através da palavra 'sozinho', o que parece ser um fenômeno inédito quando olhamos para a literatura especializada sobre modificação de eventos. Desse modo, levanto a hipótese de que as interpretações eventivas de 'sozinho' foram recentemente incorporadas no PB, uma vez que não há registros na literatura sobre essas interpretações de 'sozinho'. Assim, é esperado que as interpretações eventivas de 'sozinho' ocorram apenas em registros textuais do PB atual – o que é testado ao longo deste trabalho, sendo possível ver o resultado alcançado nos tópicos a seguir.

2. SOBRE OS MATERIAIS UTILIZADOS PARA ANÁLISE: A BASE DE DADOS

Para a realização do presente trabalho, foi utilizado como base de dados 10 textos da literatura brasileira, posteriormente divididos em 2 amostras. A primeira amostra (A)

corresponde a textos escritos antes de 1900, enquanto a segunda amostra (B) corresponde a textos escritos após 1900. A tabela a seguir apresenta os textos que compõem cada amostra.

Amostra A (Pré-1900)	Amostra B (Pós-1900)
Machado de Assis – Historias Sem Data (1884)	Júlia Lopes de Almeida – A Fallencia (1901)
Machado de Assis – Memorial de Ayres (1888)	Júlia Lopes de Almeida – Era uma vez... (1917)
Machado de Assis – Papeis Avulsos (1882)	Júlia Lopes de Almeida – A Intrusa (1908)
Machado de Assis – Poesias Completas (1864)	Bento de Villa-Moura – Nova Sapho: Tragedia Extranha (1921)
Machado de Assis – Quincas Borba (1891)	Simões Lopes Neto – Lendas do sul (1913)

Tabela 2 – Textos utilizados como base de dados

O critério de seleção dos textos para a base de dados foi: (a) todos os textos devem se localizar no contexto do português brasileiro, excluindo-se textos escritos em português europeu; (b) todos os textos devem fazer parte do gênero literário; (c) a amostra A deve ser composta por textos publicados antes de 1900 e os textos da amostra B por textos publicados após 1900. Desse modo, foram selecionados os 10 textos da tabela 2 para comporem as amostras A e B da base de dados.

3. SOBRE A INVESTIGAÇÃO: COMO OS TEXTOS FORAM ALCANÇADOS

Os textos que compõem a base de dados foram alcançados através da plataforma online Projeto Gutenberg. Trata-se de uma plataforma voluntária dedicada a digitalizar, arquivar e distribuir gratuitamente obras escritas produzidas em diferentes línguas. A escolha pelo Projeto Gutenberg se deu por diferentes critérios: (a) é uma fonte que contém obras de mais de 100 autores diferentes em português – europeu e brasileiro; (b) conveniência, uma vez que é uma plataforma de livre acesso e constantemente atualizada que pode ser facilmente acessada por qualquer usuário; (c) todos os textos na plataforma estão disponíveis no formato .txt (UTF-8), que facilita o processo de tratamento textual pela Linguagem R.

A busca dos textos na plataforma do Projeto Gutenberg foi feita da seguinte forma: (i) primeiro, acessei o catálogo de obras escritas em português; (ii) num segundo momento, conferi cada obra buscando pelas informações de data e língua, filtrando as obras que sejam escritas em português brasileiro e compreendam períodos pré e pós-1900; (iii) por fim, nas obras filtradas na segunda etapa, eu fiz uma busca pela ocorrência de 'sozinho' a partir da ferramenta de busca do navegador, para selecionar apenas as obras pré e pós-1900 em português brasileiro que apresentem ao menos uma ocorrência de 'sozinho' no texto. Após essas etapas de investigação, armazenei em um bloco de notas o endereço eletrônico para a versão .txt de cada uma das obras filtradas na plataforma do Projeto Gutenberg, conforme ilustrado na tabela 3 abaixo:

Amostra A (Pré-1900)	Amostra B (Pós-1900)
https://www.gutenberg.org/cache/epub/33056/pg33056.txt	https://www.gutenberg.org/cache/epub/69239/pg69239.txt
https://www.gutenberg.org/cache/epub/55797/pg55797.txt	https://www.gutenberg.org/cache/epub/69229/pg69229.txt
https://www.gutenberg.org/cache/epub/57001/pg57001.txt	https://www.gutenberg.org/cache/epub/67594/pg67594.txt
https://www.gutenberg.org/cache/epub/61653/pg61653.txt	https://www.gutenberg.org/cache/epub/2837/pg2837.txt
https://www.gutenberg.org/cache/epub/55682/pg55682.txt	https://www.gutenberg.org/cache/epub/26371/pg26371.txt

Tabela 3 – Endereços das versões .txt no Projeto Gutenberg das obras da base de dados

4. SOBRE O SCRIPT DESENVOLVIDO: ALGORITMO EM LINGUAGEM R

Para a automatização de tarefas de tratamento textual e extração de sentenças das amostras da base de dados selecionada, desenvolvi o script em R ilustrado na figura 1 abaixo:

```

library(stringr)
library(tokenizers)

# Diretório
setwd("/home/lisanju/Desktop/ProjetoElisa")

# Processamento do texto [ALTERE AQUI O TEXTO QUE DEVE SER PROCESSADO]
textoGutenberg <- scan(file="https://www.gutenberg.org/cache/epub/26371/pg26371.txt", what="character", sep="\n", strip.white = FALSE, blank.lines.skip = F)

# Delimitar início e fim do texto
inicio <- which(str_detect(textoGutenberg, pattern = "START OF THE PROJECT GUTENBERG EBOOK") == T) + 1
fim <- which(str_detect(textoGutenberg, pattern = "END OF THE PROJECT GUTENBERG EBOOK") == T) - 1
textoGutenberg <- textoGutenberg[inicio:fim]

# Juntar linhas que não são quebras de parágrafo
x <- textoGutenberg
y <- paste(x, collapse = " ")
z <- gsub(pattern = " ", replacement = "\n", y)
write(z, "/home/lisanju/Desktop/ProjetoElisa/textoTratado.txt")

textoTratado <- readLines("/home/lisanju/Desktop/ProjetoElisa/textoTratado.txt")

# Filtragem de palavras no texto
pal <- tokenize_words(x = textoTratado)
pal <- unlist(pal)

ocorPal <- which(str_detect(pal, "s[oó]zinh[aos]") == T)

# Filtragem de sentenças no texto
sen <- tokenize_sentences(x = textoTratado)
sen <- unlist(sen)

ocorSen <- which(str_detect(sen, "s[oó]zinh[aos]") == T)

# Transposição das sentenças para um dataframe
dataSen <- data.frame(amostra = NULL, sentenças = NULL)
dataSen <- rbind(dataSen, data.frame(amostra = "B", sentenças = sen[ocorSen]))

```

Figura 1 – Script desenvolvido em R

O script é dividido em 7 etapas, detalhadas nos subtópicos elencados a seguir.

4.1 CHAMADA DE BIBLIOTECAS

Nas primeiras linhas do script, utilizei as funções 'library(stringr)' e 'library(tokenizers)' para chamar as bibliotecas StringR e Tokenizers, que contém, respectivamente, funções úteis para o tratamento do texto e para a filtragem de sentenças.

```

library(stringr)
library(tokenizers)

```

Figura 2 – Chamada de bibliotecas

4.2 DIRETÓRIO

Através da função 'setwd()', defini o diretório de trabalho do script desenvolvido em R para uma única pasta.


```
# Diretório  
setwd("/home/lisanju/Desktop/ProjetoElisa")
```

Figura 3 – Diretório

4.3 PROCESSAMENTO DO TEXTO

Utilizei a função 'scan()' para escanear cada texto das amostras da base de dados e armazenei o resultado retornado numa variável chamada 'textoGutenberg'.

```
# Processamento do texto [ALTERE AQUI O TEXTO QUE DEVE SER PROCESSADO]  
textoGutenberg <- scan(file="https://www.gutenberg.org/cache/epub/69229/pg69229.txt", what="character", sep="\n", strip.white = FALSE, blank.lines.skip = F)
```

Figura 4 – Processamento do texto

4.4 DELIMITAR INÍCIO E FIM DO TEXTO

Para cada texto escaneado, desenvolvi no script uma etapa dedicada a delimitar o início e o fim do conteúdo textual – excluindo, por exemplo, as notas de observações do Projeto Gutenberg, padronizadas e presentes em todas as obras arquivadas no site da plataforma.

```
# Delimitar início e fim do texto  
inicio <- which(str_detect(textoGutenberg, pattern = "START OF THE PROJECT GUTENBERG EBOOK") == T) + 1  
fim <- which(str_detect(textoGutenberg, pattern = "END OF THE PROJECT GUTENBERG EBOOK") == T) - 1  
textoGutenberg <- textoGutenberg[inicio:fim]
```

Figura 5 – Delimitar início e fim do texto

Para fazer essa delimitação, busquei pelos padrões 'START OF THE PROJECT GUTENBERG EBOOK' e 'END OF THE PROJECT GUTENBERG EBOOK', que se repetem no arquivo .txt de todas as obras coletadas.

4.5 JUNTAR LINHAS QUE NÃO SÃO QUEBRAS DE PARÁGRAFO

Para facilitar na etapa posterior de extração de palavras e sentenças, fiz um tratamento textual nas obras escaneadas para juntar linhas que não são quebras de parágrafo.

```
# Juntar linhas que não são quebras de parágrafo
x <- textoGutenberg
y <- paste(x, collapse = " ")
z <- gsub(pattern = " ", replacement = "\n", y)
write(z, "/home/lisanju/Desktop/ProjetoElisa/textoTratado.txt")

textoTratado <- readLines("/home/lisanju/Desktop/ProjetoElisa/textoTratado.txt")
```

Figura 6 – Juntar linhas que não são quebras de parágrafo

Após o tratamento, escrevo um novo arquivo tratado e abro ele através da função 'readLines()', armazenando seu resultado numa variável chamada 'textoTratado'.

4.6 FILTRAGEM DE PALAVRAS E SENTENÇAS NO TEXTO

Fiz uma filtragem através da biblioteca Tokenizers das ocorrências da palavra 'sozinho' nos textos escaneados e armazenei o resultado na variável 'ocorPal'. Também fiz o mesmo processo de filtragem para sentenças que são compostas pela palavra 'sozinho', armazenando o resultado numa variável 'ocorSen'.

```
# Filtragem de palavras no texto
pal <- tokenize_words(x = textoTratado)
pal <- unlist(pal)

ocorPal <- which(str_detect(pal, "s[oóô]zinh[aos]") == T)

# Filtragem de sentenças no texto
sen <- tokenize_sentences(x = textoTratado)
sen <- unlist(sen)

ocorSen <- which(str_detect(sen, "s[oóô]zinh[aos]") == T)
```

Figura 7 – Filtragem de palavras e sentenças no texto

4.7 TRANSPOSIÇÃO DAS SENTENÇAS PARA UM DATAFRAME

Por fim, a última etapa do script desenvolvido em R é a de armazenar os resultados obtidos na filtragem de sentenças em um dataframe chamado 'dataSen'. Isso facilitará na etapa manual de rotular cada uma das interpretações ativadas por 'sozinho' nas sentenças filtradas. No dataframe, há uma coluna para indicar se a sentença faz parte da Amostra A ou B e outra coluna contendo a sentença de fato.

```
# Transposição das sentenças para um dataframe
dataSen <- data.frame(amostra = NULL, sentenças = NULL)
dataSen <- rbind(dataSen, data.frame(amostra = "B", sentenças = sen[ocorSen]))
```

Figura 8 – Transposição das sentenças para um dataframe

4.8 ROTULAÇÃO MANUAL DAS INTERPRETAÇÕES DE 'SOZINHO'

Com o dataframe gerado pelo script em R após o tratamento textual e da filtragem das sentenças das amostras, manualmente rotulei cada uma das interpretações ativadas por 'sozinho' através dos rótulos a seguir:

- [1] espaço-temporal
- [2] mereológica
- [3] modificador argumental
- [4] emocional
- [5] comportamental
- [6] (anti-)assistida
- [7] (anti-)causal

Os rótulos correspondem com a descrição feita em Basso & Rodrigues (2024) e capturam as diferentes interpretações de 'sozinho' a partir de seu escopo na sentença. Como apresentado anteriormente, as interpretações rotuladas como [6] e [7] se tratam das interpretações eventivas de 'sozinho', que é o fenômeno estudado neste trabalho. Conforme minha hipótese principal, as interpretações eventivas não devem ocorrer em textos da amostra A (pré-1900), mas sim em textos da amostra B (pós-1900), uma vez que é uma interpretação recentemente

incorporada pelo 'sozinho' – logo, deve aparecer apenas em textos de um português brasileiro atual. No tópico a seguir, discuto os resultados obtidos no trabalho e as análises estatísticas realizadas.

5. ANÁLISE ESTATÍSTICA E RESULTADOS

Para realizar a análise estatística dos resultados, desenvolvi um script à parte em R para a geração de gráficos a partir da planilha rotulada manualmente com as interpretações de 'sozinho'.

```
# Leitura da planilha rotulada
arqRotulado <- read.table("/home/lisanju/Desktop/ProjetoElisa/planilha.txt", header = TRUE, sep = "\t", fill = TRUE, stringsAsFactors = FALSE, quote = "")

# Gráficos
freqTable <- table(arqRotulado$amostra, arqRotulado$inter)
barplot(freqTable, beside = FALSE, col = c("blue", "red"),
        legend.text = TRUE, args.legend = list(title = "Amostras"),
        xlab = "Interpretações",
        ylab = "Ocorrências",
        main = "Quantificação de 'sozinho' em A e B")
```

Figura 9 – Geração de gráficos

O script abre o arquivo .txt contendo as sentenças rotuladas como uma table e, então, armazena a frequência dos valores do arquivo. Assim, é gerado um gráfico barplot empilhado a partir da função 'barplot()' para a visualização do resultado da quantificação:

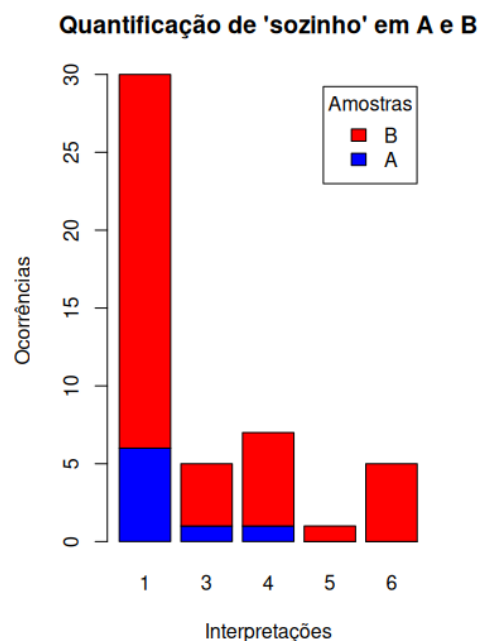


Gráfico 1 – Quantificação de 'sozinho' em A e B

A escolha pelo gráfico barplot empilhado se deu pelo fato de ser o gráfico que captura melhor o resultado da quantificação, sem uma poluição com várias barras ilustradas e sendo intuitivo para capturar rapidamente as informações necessárias.

Além disso, também quantifiquei o número de ocorrências da palavra 'sozinho' pela quantidade de palavras de cada obra das amostras A e B através da função 'tokenize_words()':

```
# Filtragem de palavras no texto
pal <- tokenize_words(x = textoTratado)
pal <- unlist(pal)

ocorPal <- which(str_detect(pal, "s[oóô]zinh[aos]") == T)
```

Figura 10 – Quantificação de palavras e ocorrências de 'sozinho'

Amostra A) Machado de Assis – Historias Sem Data (1884)

1 ocorrência / 51638 palavras

Amostra A) Machado de Assis – Memorial de Ayres (1888)

4 ocorrências / 51338 palavras

Amostra A) Machado de Assis – Papeis Avulsos (1882)

1 ocorrência / 55305 palavras

Amostra A) Machado de Assis – Poesias Completas (1864)

1 ocorrência / 44948 palavras

Amostra A) Machado de Assis – Quincas Borba (1891)

1 ocorrência / 78427 palavras

Amostra B) Júlia Lopes de Almeida – A Falência (1901)

1 ocorrência / 6410 palavras

Amostra B) Júlia Lopes de Almeida – Era uma vez... (1917)

15 ocorrências / 74005 palavras

Amostra B) Júlia Lopes de Almeida – A Intrusa (1908)

24 ocorrências / 55487 palavras

Amostra B) Bento de Villa-Moura – Nova Sapho: Tragedia Extranha (1921)

6 ocorrências / 18781 palavras

Amostra B) Simões Lopes Neto – Lendas do sul (1913)

1 ocorrência / 46774 palavras

Desse modo, temos que, para a os textos da amostra A, há um total de 281656 palavras, em que há 8 ocorrências da palavra 'sozinho'. Para os textos da amostra B, há um total de 201457 palavras, em que há 47 ocorrências da palavra 'sozinho'.

5.1 DISCUSSÃO DOS RESULTADOS

A partir do gráfico feito com base nas amostras utilizadas especificamente neste trabalho, é possível observar que não há ocorrências de interpretações eventivas nos textos da amostra A. É possível afirmar, também, que tanto na amostra A quanto na amostra B a interpretação espaço-temporal de 'sozinho' (rótulo 1) se mantém como a interpretação mais frequente. É possível explicar esse comportamento se considerarmos que as diferentes interpretações de 'sozinho' possuem um núcleo de significado em comum, em que esse núcleo parte da noção concreta de não possuir companhia num espaço-tempo e, por um processo de abstração, desdobra-se em interpretações mais abstratas, como as eventivas e de modificação argumental.

Outro comportamento que se pode observar pelo gráfico é que todas as interpretações ocorridas na amostra A também ocorrem na amostra B, mas o contrário não. As interpretações comportamental (rótulo 5) e (anti-)assistida (rótulo 6) não ocorrem na amostra A, apesar de ocorrerem na amostra B. A partir disso, pode-se concluir que ao longo do tempo, da amostra A para a amostra B, as interpretações de 'sozinho' que já eram utilizadas se mantiveram e, além disso, 'sozinho' passou a ser usado para outras interpretações também.

Dado os resultados obtidos a partir da quantificação de ocorrências da palavra 'sozinho' pela número de palavras de cada obra, uma conclusão possível de se alcançar é a de que há um fator de influência de estilo de escrita para a ocorrência ou não das interpretações de 'sozinho' – pode-se ver, por exemplo, que as obras de Júlia Lopes de Almeida possuem 5 vezes mais ocorrências de 'sozinho' do que as obras de Machado de Assis. Ainda sobre a quantificação de palavras, nota-se também que, mesmo em menor número de palavras total, a

amostra B possui bem mais ocorrências de 'sozinho' do que a amostra A – sendo este outro motivo para concluir que a presença de uma interpretação eventiva de 'sozinho' está relacionada com o estilo do autor.

REFERÊNCIA

BASSO, Renato; RODRIGUES, Elisa. **Sobre a semântica de ‘sozinho’: uma descrição de suas interpretações.** Cadernos de Estudos Linguísticos da Unicamp, Campinas. 2024