



Universidade Estadual de Feira De Santana
Programa de Pós-Graduação em Computação Aplicada

SÍNTESE COMPUTACIONAL DE POEMAS BASEADOS EM PROSA LITERÁRIA BRASILEIRA

Ana Cleyge Silva de Azevedo

Feira de Santana

2018



Universidade Estadual de Feira De Santana
Programa de Pós-Graduação em Computação Aplicada

Ana Cleyge Silva de Azevedo

Síntese Computacional de Poemas Baseados em Prosa Literária Brasileira

Dissertação apresentada à Universidade Estadual de Feira de Santana como parte dos requisitos para a obtenção do título de Mestre em Computação Aplicada.

Orientador: Prof. Angelo Loula (UEFS)

Coorientador: Prof. João Queiroz (UFJF)

Feira de Santana

2018

Ficha Catalográfica – Biblioteca Central Julieta Carteado

A986s Azevedo, Ana Cleyge Silva de
 Síntese computacional de poemas baseados em prosa literaria
 brasileira./ Ana Cleyge Silva de Azevedo. – 2018.
 90f.: il.

 Orientador: Angelo Loula
 Coorientador: João Queiroz
 Dissertação (mestrado) – Universidade Estadual de Feira de Santana,
 Programa de Pós-Graduação em Computação Aplicada, 2018.

 1.Algoritmo. 2.Poemas. 3.Sistema MIVES. 4.PROPOE. I.Loula,
 Angelo, orient. II.Queiroz, João, coorient. III.Universidade Estadual
 de Feira de Santana. IV. Título.

 CDU: 004.421

 Maria de Fátima de Jesus Moreira – Bibliotecária – CRB5/1120

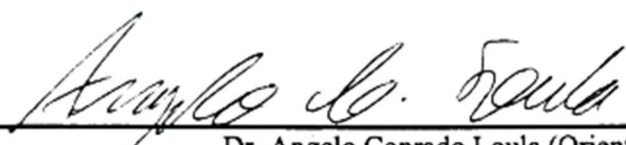
Ana Cleyge Silva de Azevedo

**Síntese Computacional de Poemas Baseados em Prosa Literária
Brasileira**

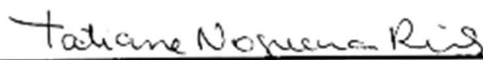
Dissertação apresentada à Universidade Estadual de Feira de Santana como parte dos requisitos para a obtenção do título de Mestre em Computação Aplicada.

Feira de Santana, 06 de setembro de 2018

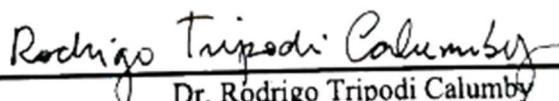
BANCA EXAMINADORA



Dr. Angelo Conrado Loula (Orientador)
Universidade Estadual de Feira de Santana



Dra. Tatiane Nogueira Rios
Universidade Federal da Bahia



Dr. Rodrigo Tripodi Calumby
Universidade Estadual de Feira de Santana

Abstract

The computational generation of poems has been studied internationally for decades being improved over time. The most recent contributions consider multiple linguistic levels in the generation of poems, particularly phonological parallels, metrical structures of versification, rhythmic patterns and rhyme. On the other side, the presence of metric structure was identified in literary prose and the computational system MIVES (Mining Verse Structure) is able to automatically find such structures in prose. Therefore, this work presents the PROPOE (Prose to Poem Automatic Generator), a system capable of receiving the metric sentence data from MIVES to generate poems, filtering sentences to build poems. PROPOE is implemented based on the greedy algorithm, capable of performing combinatorial optimization of verses in stanzas, considering rhythmic aspects such as rhyme and tonic positioning. The poems obtained are evaluated by the system considering multiple criteria to obtain a general score. Considering several prose works as a source of sentences, the results achieved by PROPOE are presented and analyzed in multiple examples.

Keywords: poem, rhythmic patterns, rhyme, algorithm, computing.

Resumo

A geração computacional de poemas tem sido estudada internacionalmente por décadas sendo aprimorada ao longo do tempo. As contribuições mais recentes consideram múltiplos níveis linguísticos na geração de poemas, particularmente paralelismos fonológicos, estruturas métricas de versificação, padrões rítmicos e rima. Por outro lado, a presença de estruturas métricas foi identificada em prosa literária e o sistema computacional MIVES é capaz de encontrar de forma automatizada em obras em prosa. Assim, este trabalho apresenta o sistema PROPOE (Prose to Poem Automatic Generator) capaz de receber os dados de sentenças métricas advindas do MIVES para gerar poemas, filtrando sentenças para montagem do poema. O PROPOE é implementado com base em um algoritmo guloso, capaz de realizar otimização combinatória dos versos em estrofes, considerando aspectos rítmicos como rima e posicionamento tônico. Os poemas obtidos são avaliados pelo sistema considerando múltiplos critérios para obter uma pontuação geral. A partir de variadas obras em prosa como fonte de sentenças, os resultados alcançados pelo PROPOE são apresentados e analisados em múltiplos exemplos.

Palavras-chave: poema, padrões rítmicos, rima, algoritmo, computação.

Prefácio

Esta dissertação de mestrado foi submetida a Universidade Estadual de Feira de Santana (UEFS) como requisito parcial para obtenção do grau de Mestre em Computação Aplicada.

A dissertação foi desenvolvida dentro do Programa de Pós-Graduação em Computação Aplicada (PGCA) tendo como orientador o Dr. **Angelo Loula** e co-orientador o Dr. **João Queiroz**.

Agradecimentos

Ao meu orientador, Professor Dr. Angelo Loula, que me incentivou a enfrentar e vencer os desafios ao longo do curso, sempre disponível a ajudar de orientar nas tomadas de decisões.

Ao meu co-orientador, Professor Dr. João Queiroz, que mesmo à distância, realizou intervenções significativas na elaboração do trabalho.

A toda a minha família e amigos que compreenderam a minha ausência em momentos singulares das nossas vidas.

Ao meu esposo, Ciro José Azevedo, que contribui diretamente para minha realização profissional e acadêmica. Provendo todos os recursos necessários, atendendo às solicitações, sempre atencioso, dedicado. Um grande companheiro.

Aos meus filhos, Laura, Ciro Filho e Fernanda, como participantes ativos de todo o processo. Amor incondicional.

Sumário

Abstract	i
Resumo	ii
Prefácio	iii
Agradecimentos	iv
Sumário	x
Lista de Tabelas	xii
Lista de Figuras	xiv
1. Introdução	1
1.1 <i>Objetivos</i>	4
1.2 <i>Organização do Trabalho</i>	4
2. Fundamentação Teórica	6
2.1 <i>Poema</i>	6
2.1.1 <i>Métrica</i>	7
2.1.2 <i>Ritmo</i>	9
2.1.3 <i>Verso</i>	10
2.1.4 <i>Estrofe</i>	11
2.1.5 <i>Rima</i>	11
2.2 <i>Estrutura Métrica em Prosa Literária Brasileira</i>	15
2.3 <i>Fundamentação Computacional</i>	16
2.3.1 <i>Idiomas</i>	16
2.3.2 <i>Dados de Entrada no Sistema</i>	18
2.3.3 <i>Técnicas Aplicadas</i>	21
2.3.3.1 <i>Raciocínio Baseado em Casos</i>	21
2.3.3.2 <i>Tabela de Geração</i>	23
2.3.3.3 <i>Modelo Estocástico Hill Climbing</i>	25
2.3.3.4 <i>Algoritmo Evolutivo</i>	27
2.3.3.5 <i>Quadro Negro</i>	28
2.3.3.6 <i>Rede Neural</i>	28
2.3.4 <i>Avaliações dos Geradores de Poema</i>	30
3. O Sistema	34
3.1 <i>Texto com estruturas versificadas</i>	35
3.2 <i>Organização de Sentenças por Sílabas</i>	36

3.3 Filtro das Sentenças.....	38
3.4 Montagem do poema.....	38
3.5 Avaliação do poema.....	44
3.6 Interface do Sistema.....	46
4. Resultados Alcançados	50
4.1 Os Sertões	50
4.2 Recordações do Escrivão Isaiás Caminha	56
4.3 Contrastes e Confrontos	59
4.4 Os Lusíadas.....	62
5. Considerações Finais	66
6. Referências Bibliográficas	68
7. Anexo 01	71

Lista de Tabelas

Tabela 2.1: Sentença Escandida do Poema "Remorso" de Olavo Bilac.....	8
Tabela 2.2: Trecho do poema "Remorso" de Olavo Bilac.	9
Tabela 2.3: Trecho do poema "Uruguai" de Basílio da Gama.	10
Tabela 2.4: Trecho do poema "Paisagem" de Vinícius de Moraes.	10
Tabela 2.5: Trecho do poema Fernando Pessoa.	11
Tabela 2.6: Rima coroadada.....	12
Tabela 2.7: Rima leonina.....	12
Tabela 2.8: Rima com eco.	12
Tabela 2.9: Rima consoante.....	12
Tabela 2.10: Rima toante.....	13
Tabela 2.11: Rima Cruzada.	13
Tabela 2.12: Rima Emparelhada.	13
Tabela 2.13: Rima Interpolada.	14
Tabela 2.14: Rima Mista.	14
Tabela 2.15: Rima Perfeita.	14
Tabela 2.16: Rima por Acentuação.	15
Tabela 2.17: Exemplos de Entrada na Tabela "the dog saw the cat" [Chrismartin, Tobing e Manurung, 2015].	24
Tabela 2.18: Comparação entre trabalhos dos autores.	33
Tabela 3.1: Estruturas versificadas extraídos do MIVES.....	36
Tabela 3.2: Sentenças escandidas do MIVES.	36
Tabela 3.3: Algumas sentenças do arquivo XML.	37
Tabela 3.4: Quantidade de sílabas no poema.	40
Tabela 3.5: Rima mista.....	41
Tabela 3.6: Posicionamento dos versos.....	41
Tabela 3.7: Poemas com diferentes combinações de versos.	42
Tabela 3.8: Poemas com diferentes combinações de versos.	43
Tabela 3.9: Avaliação do esquema rítmico.	44
Tabela 3.10: Avaliação das Tônicas Iguais.	45
Tabela 3.11: Avaliação da similaridade por acentuação.	45
Tabela 3.12: Avaliação da similaridade fonética do poema.....	46
Tabela 4.1: Frequência de Algumas Sentenças por Sílabas.....	51
Tabela 4.2: Poema de Métrica Heptassílaba.....	54
Tabela 4.3: Poema de Métrica Decassílaba.	55
Tabela 4.4: Poema de Métrica Decassílaba.	55
Tabela 4.5: Poema de Métrica Hexassílaba.....	58
Tabela 4.6: Poema de Métrica Bárbara.	58
Tabela 4.7: Poema de Versos Heterométricos.....	59

Tabela 4.8: Poema Hexassílabo.....	61
Tabela 4.9: Poema Heptassílabo.....	61
Tabela 4.10: Poema Heterométrico.	62
Tabela 4.11: Poema Decassílabo.....	64
Tabela 4.12: Poema Hendecassílabo.	65
Tabela 4.13: Poema Dodecassílabo.....	65

Lista de Figuras

Figura 2.1: Geração de versos usando RBC. Adaptada de Gervás [2001].	22
Figura 2.2: Arquiteutra do PoeTryMe [Oliveira, 2012].	25
Figura 2.3: Diagrama idealizado de uma pesquisa estocástica [Manurung, Ritchie e Thompson, 2000].	26
Figura 2.4: Os três processos Reiter em uma arquitetura de pipeline [Manurung, 2004].	27
Figura 2.5: Arquitetura do quadro negro usada no sistema de Misztal e Indurkhya [2014].	28
Figura 2.6: Visão geral do Hefaz [Ghazvininejad <i>et. al.</i> , 2017].	30
Figura 3.1: Etapas de processo do sistema.	35
Figura 3.2: Montagem do Poema.	39
Figura 3.3: Interface de Visualização de Análise dos Resultados.	47
Figura 3.4: Poema exibido em forma de verso e de sentença escandida.	48
Figura 3.5: Interface de Visualização para Análise dos Resultados.	49
Figura 4.1: Frequência de Algumas Sentenças por Sílabas.	51
Figura 4.2: Frequência de Sentenças Decassílabas por Sílaba.	52
Figura 4.3: Frequência de Sentenças Dodecassílabas por Sílaba.	52
Figura 4.4: Frequência de Sentenças Decassílabas por Esquema Rítmico.	53
Figura 4.5: Frequência de Sentenças Dodecassílabas por Esquema Rítmico.	53
Figura 4.6: Frequência de Sentenças por Sílaba.	56
Figura 4.7: Frequência de Sentenças Hexassílabas por Métrica.	57
Figura 4.8: Frequência de Sentenças Octossílabas por Métrica.	57
Figura 4.9: Frequência de Sentenças por Sílaba.	60
Figura 4.10: Frequência de Sentenças Heptassílabas por Métrica.	60
Figura 4.11: Frequência de Sentenças Hexassílabas por Métrica.	60
Figura 4.12: Frequência de Sentenças por Sílabas.	63
Figura 4.13: Frequência de Sentenças Decassílabas por Métrica.	63
Figura 4.14: Frequência de Sentenças Hendecassílabas por Métrica.	64

Capítulo 1

Introdução

“Por vezes sentimos que aquilo que fazemos não é senão uma gota de água no mar. Mas o mar seria menor se lhe faltasse uma gota.”

-- Madre Tereza de Calcutá

A composição estrutural de um poema costuma ser realizada com base na aplicação do conhecimento das regras gramaticais de paralelismos fonológicos, léxicos e semânticos, responsáveis por atribuir uma descrição estrutural à sentença. Tais regras estabelecidas pela gramática, são responsáveis por garantir que o poeta manipule e organize palavras, oferecendo uma percepção de semelhança entre sons, ritmo, e que as suas combinações possam fazer sentido. Neste contexto, Goldstein [2006] considera que o ritmo em um poema está diretamente ligado às noções de metro de um verso da tradição literária.

Um verso é separado em sílabas poéticas, e é feita a identificação do número de sílabas, o seu posicionamento e as suas combinações, bem como a marcação e regularidade das sílabas fortes. A tarefa de separação de sílabas de um verso é conhecida como escanção de versos, e as sílabas de um mesmo verso podem ser caracterizadas com mais de uma métrica, devido a escolha do autor por realizar determinadas junções ou separações silábicas, justificando o aumento ou diminuição da sua métrica.

A partir da métrica também é possível perceber a rima, conhecida por uma sequência de sons, que pode aparecer no poema de forma isolada, e por isso classificada como rima órfã. Nas suas demais aparições, ela costuma ocorrer minimamente aos pares, fazendo correspondência pela semelhança sonora entre as sílabas poéticas de palavras. Dessa forma, a rima entre palavras pode ser marcada pela coincidência de sons entre as últimas sílabas poéticas das mesmas em final de versos, ou mesmo em palavras que rimam no seu interior. Além da posição, a rima pode ser classificada pela distribuição entre os versos, pela categoria gramatical e pela extensão dos seus sons. São estes os aspectos formais referentes à estrutura externa de um poema.

Um tipo de texto que algumas vezes contrasta com a estrutura externa de um poema é a prosa. Este tipo de texto, pode não conter aspectos de rimas ou ritmos regulares, pois em geral, a prosa apresenta narração e linguagem escrita de forma contínua, com marcação em parágrafos. Desta forma, D’Onofrio [2007] afirma que a diferença entre estas obras literárias, é a presença maior de elementos fônicos, lexicais, sintáticos e semânticos, constitutivos da linguagem poética.

Apesar desta distinção entre os textos, certos padrões comuns de metrificação foram observados por Almeida [1946] na prosa literária d'Os Sertões escrita por Euclides da Cunha. A partir dos versos identificados, Almeida [1946] considerou o escritor dessa obra como um mestre na prática da difícil arte de produzir versos livres. Após cinquenta anos, Campos [1996] apresentou um estudo realizado sobre os aspectos poéticos na prosa literária d'Os Sertões, quando encontrou metrificações silábicas de muitos decassílabos e dodecassílabos, com características poéticas, sendo que muitos deles possuem marcações regulares em suas sílabas tônicas.

As estruturas classificadas como versos decassílabos, e dodecassílabos, são representadas por versos de 10 e 12 sílabas respectivamente, com marcação frequente de sílaba tônica nas mesmas posições entre alguns dos versos. Além disso, Almeida [1946] reconhece que certas passagens da prosa poderiam ser escritas em forma de poema, devido a identificação em fragmentos da obra de Os Sertões de várias estruturas métricas marcantes e bem definidas em final de parágrafos. Tais métricas são identificadas com a separação em sílabas poéticas e a contagem das mesmas, sendo esta atividade conhecida como escanção. Dessa forma, é apresentado como exemplo, um dos poemas que Almeida [1946] identificou partir de sentenças extraídas da prosa literária:

A vaquejada

De repente estruge ao lado
um estríduo tropel de cascos sobre pedras,
um estrépido de galhos estalando,
um estalar de chifres embatendo;
tufa nos ares, em novelos,
uma nuvem de pó;
rompe, a súbitas, na clareira,
embolada, uma ponta de gado;
e, logo após,
sobre o cavalo que estaca esbarrado,
o vaqueiro, teso nos estribos... (p. 126).

Considerando a possibilidade de automatizar a tarefa de identificação métrica em texto em prosa, Carvalho, Loula e Queiroz [2016] desenvolveram o sistema computacional MIVES (*Mining Verse Structure*). Este sistema tem como saída um conjunto de segmentos frásicos com estruturas métricas, e suas classificações relacionadas a sentenças escandidas, número de sílabas poéticas e marcação das sílabas tônicas. Porém, a composição automatizada de poemas a partir destas sentenças é uma oportunidade ainda em aberto, que requer um estudo relacionado a distribuição gráfica do poema, para que os seus versos nas estrofes sejam estruturados com base em aspectos de rima e ritmo. Assim, as sentenças versificadas do MIVES são exploradas neste trabalho, como material para implementação do sistema PROPOE (*Prose to Poem Automatic Generator*).

A geração automática de poemas tem sido um tópico de pesquisa de longa data na computação. As sistemáticas tentativas de composição computacional de poesia verbal são feitas desde 1959,

em diferentes idiomas, e os desenvolvimentos ocorridos para esta finalidade são decorrentes do escopo histórico do fenômeno conhecido como “arte algorítmica” ou “arte computacional” [Antonio, 2016]. Estes algoritmos têm como funcionalidade combinar palavras para gerar frases consideradas poéticas como as de Lutz [1959] por apresentar imagens formadas de palavras em forma de geometria espacial como as de Souza *apud* Antonio [2016], ou misturar frases de crônicas e poesias para gerar novos poemas no sistema de Balestrini *apud* Antonio [2016].

Além da arte computacional, uma outra maneira de gerar poemas automáticos é por meio da criatividade computacional, pois de acordo com Boden [2009] essa abordagem viabiliza tanto a capacidade de avanço de artefatos novos, valiosos e surpreendentes relacionados à produção intelectual concreta, quanto às ideias consideradas produções intelectuais abstratas. Sendo a criatividade computacional um dos ramos da inteligência artificial, através de técnicas computacionais, é possível realizar atividades criativas. Porém, Boden [2009] afirma que é preciso garantir que os poemas sejam criados com base nos padrões aceitáveis pela norma literária, além de serem compreensíveis.

A criatividade computacional foi explorada por alguns pesquisadores para gerar poemas. Tais pesquisas foram iniciadas por Gervás [2000], Manurung Ritchie e Thompson [2000] e Oliveira [2009] e têm aprimorado as suas implementações ao longo de todos esses anos, com uso de variadas técnicas, desenvolvendo sistemas capazes de gerar poemas com formas e conteúdos variados. Além disso, a Microsoft publicou um livro na China com 139 poemas gerados por um computador, como fruto do método desenvolvido por Zhou [2008] para criar dísticos clássicos chineses.

Dentre as pesquisas realizadas sobre os sistemas que fazem a geração de poemas, foram encontradas contribuições em variados idiomas, como alemão, castelhano, chinês, espanhol, indonésio, inglês, português europeu, cujos propósitos assemelham-se, diferindo-se pela utilização de variados métodos, metodologias ou técnicas. Desde o primeiro *software* desenvolvido, portanto, variadas abordagens foram aplicadas, e torna-se imprescindível uma análise geral delas. Nos trabalhos analisados, foram identificados aspectos relacionados à implementação do *software*, como escolha da fonte a ser explorada, que pode ser em variados tipos de textos, os requisitos estabelecidos para implementação do sistema; o método aplicado como forma de organizar os processos; e a avaliação dos resultados alcançados.

Com o passar dos anos, foram feitas algumas outras produções sobre o tema, explorando novos métodos. A comunidade científica também foi ampliando o interesse no tema, principalmente por pesquisadores do domínio da inteligência artificial, a exemplo do *software* desenvolvido por Gervás [2000] que gera poemas a partir de obras literárias, permitindo a alteração com acréscimo de palavras ou troca das mesmas, consideradas adequadas ao contexto poético, Oliveira [2012] e por Ghazvininejad [2016] para gerar poemas com combinação entre palavras sinônimas; o gerador de poemas de Zhang e Lapata [2014] que faz análise de repositório de poemas online para gerar novos poemas.

Além da escolha pela técnica mais adequada para desenvolver um gerador de poemas que possa ter um bom desempenho na sua atividade de processamento, cada algoritmo desenvolvido é construído com base em regras de combinação frásica, padrões de versos, formas do poema, e verificação do significado das palavras usadas. Mas em sua maioria, são eleitas algumas dessas regras nos geradores de poemas, pois nem todas as produções poéticas necessitam de extrema precisão [Oliveira, 2017]. Para isso, alguns dos geradores podem transmitir uma mensagem com pouca clareza por priorizar um ritmo agradável, em uma métrica regular.

A qualidade de um poema é considerada por Gervás [2016] como a existência do equilíbrio entre a forma e o conteúdo, visto que o conteúdo não pode surgir aleatoriamente a partir do resultado da exploração de um determinado espaço conceitual em busca de um valor estético. Por isso, existem dois níveis de limitação na geração de poemas, um deles é referente a relação em que as palavras mantêm no texto, pois estas precisam transmitir uma mensagem. O outro nível diz respeito a definição da forma poética na distribuição dos versos no poema, seja pelo número de linhas de comprimentos específicos de sílabas, ou pelas palavras que rimam no final dos versos.

Um dos fatos relevantes observado, e que impulsiona o andamento desta pesquisa, corresponde a possibilidade de realizar combinações entre um grande número de sentenças. Para isso, é preciso fazer a identificação de semelhanças entre os esquemas rítmicos, explorando diversos tipos de rimas, com um grande número de sílabas, e tudo isso combinado à formas diferentes de busca por similaridades entre versos, considerando esquema rítmico ou posicionamento de sílabas tônicas.

A qualidade do poema gerado também deve ser analisada pelo próprio sistema, por meio de uma avaliação realizada através de pontuação sob diferentes perspectivas relacionadas a fonética, esquema rítmico e similaridade entre as sílabas tônicas dos versos. Apesar da ausência de avaliação para aspectos semânticos, os poemas gerados farão sentido por serem extraídos de uma mesma obra literária.

Este trabalho apresenta um sistema que promove uma nova forma de gerar poemas automáticos, por meio de um algoritmo capaz de solucionar problemas de otimização combinatória. Dessa forma, são consideradas restrições específicas de conteúdo, e exploração de aspectos rítmicos e de rima, a partir de sentenças versificadas, extraídas da prosa literária. Com isso, é exibido o reflexo das impressões do autor, e de seus sentimentos, através de um estilo reduzido de suas estruturas frásicas.

1.1 Objetivos

O objetivo geral deste trabalho é propor um sistema computacional capaz de organizar sentenças metrificadas de estruturas de versificação da prosa literária brasileira, para gerar poemas, com base em critérios e padrões rítmicos estabelecidas para esta finalidade. Deste objetivo, surgem como objetivos específicos:

- Desenvolver um sistema computacional com otimização combinatória;
- Estabelecer critérios de otimização combinatória de rima e ritmo;
- Implementar uma interface que possa filtrar os dados para compor o poema;
- Explorar variadas obras literárias através do sistema implementado;
- Avaliar as produções realizadas pelo sistema conforme critérios definidos;
- Disponibilizar um novo conjunto de *corpora* poéticos a serem utilizados em trabalhos futuros.

1.2 Organização do Trabalho

Esta pesquisa está disposta em cinco capítulos, já iniciado pela introdução, na qual o leitor pode obter um prévio relato daquilo que será abordado nos demais capítulos.

No Capítulo 2 é feita a abordagem da fundamentação teórica, contendo uma revisão dos artigos estudados sobre trabalhos relacionados, bem como o material referente à revisão da literatura, com base nos conceitos fundamentais para compreensão da ferramenta desenvolvida.

O Capítulo 3 apresenta a metodologia, descrevendo a sistematização dos processos para implementação do sistema, bem como os instrumentos empregados para pesquisa, avaliando as capacidades e limitações do sistema.

O Capítulo 4 é composto por resultados alcançados durante o filtro de dados, como também no processo de otimização combinatória dos mesmos para geração de poemas, tendo por base os objetivos alcançados com a prática do que foi estabelecido no capítulo da metodologia.

No Capítulo 5 são feitas as considerações finais sobre todo trabalho de pesquisa, tendo uma correspondência entre os objetivos propostos e os resultados alcançados, com indicativos que possam viabilizar a continuidade desta pesquisa.

Capítulo 2

Fundamentação Teórica

“Se enxerguei mais longe, foi porque estava sobre os ombros de gigantes.”

-- Isaac Newton

Neste capítulo são apresentados elementos correspondentes à literatura brasileira, e suas definições, que são determinadas pelas normas gramaticais para estrutura de um poema, especificamente com ênfase em rima e ritmo, bem como os seus aspectos formais. Estes são assuntos contemplados no sistema proposto, descrevendo-os com exemplos práticos, desde a contagem de sílabas poéticas até a montagem das estrofes, complementando com aspectos lexicais, sintáticos e semânticos.

Na sequência é apresentado um esboço referente às considerações importantes a serem feitas sobre aspectos computacionais. E por fim, é feita uma abordagem sobre os sistemas desenvolvidos desde os primórdios da computação digital, enfatizando as técnicas mais novas utilizadas na geração de poemas, e as experiências vividas pelos pesquisadores. Com isso, o leitor é informado sobre a evolução tecnológica ocorrida ao longo dos anos na geração automática de poemas.

2.1 Poema

Apesar da poesia e poema estarem interligadas, ambos não são sinônimos, e Silva [2014] afirma que em todo poema tem poesia, mas esta não está presente apenas em poemas. A poesia está presente nas expressões artísticas do ser criativo, capazes de sugerir emoções através da linguagem que combina sons, ritmos e significados, estando presente em uma música, no teatro, em objetos, e em versos exibidos por um poema, que pode caracterizar e provocar variados sentidos ao receptor da mensagem, através de efeitos de sons, ritmo, rima, aliteração, metáforas e personificações.

Seguindo a mesma linha de pensamento, Goldstein [2006] considera a poesia como algo que transcende o texto, pois está presente em variadas obras artísticas, que são consideradas poéticas, por despertar no leitor um efeito marcante do poema, que convida à releitura, e permite que seja feita mais de uma interpretação. Assim, são abordados posteriormente os aspectos formais que um poema possui, e sobre o que deve ser analisado em um texto em forma de prosa para identificação de traços poéticos.

A formação de um poema pode ou não obedecer a regras clássicas estabelecidas pela métrica, ou seja, pelo posicionamento de sílabas acentuadas em cada verso, com rimas marcadas regularmente, exibindo semelhança fônica no final de certas palavras, seja ao término de cada sentença ou no meio delas, de forma simétrica. Porém, de acordo com Goldstein [2006] é possível que nada disso aconteça, quando a regra é estabelecida com base na desconstrução de regras, ou do pensamento livre.

De acordo com Proença [2007] no final do século XIX surge um novo tipo de verso, classificado como verso livre, tendo como característica a valorização da entonação, das pausas, e da velocidade da enunciação no contexto do poema, ou seja, seu ritmo é apoiado na combinação entre entoação e pausa. Além disso, Proença [2007] enfatiza que a entoação está relacionada a melodia que caracteriza o enunciado, sendo alternada pela escala de elevação da voz.

A articulação entre a seleção e a combinação de palavras, além de sinais decorrentes de paralelismo, como aspectos rítmico, sintático e sonoro, são consideradas por Goldstein [2006] como pistas para que o leitor possa identificar traços que caracterizam o poema. Tais aspectos são caracterizados pelas relações de semelhanças estabelecidas entre as palavras, ou mesmo entre os versos, para completar o sentido no poema.

Alguns poemas possuem padrão fixo em sua estrutura, e são facilmente identificados pelos seus versos, em ode, canção, quadrinha, soneto, haicai, dentre outros, mas na forma contemporânea, apresenta mudanças quanto aos estilos literários, flexibilizando a liberdade de expressão, através de versos associados a imagens, de poemas que requerem complemento a ser feito pela reflexão do leitor.

Um poema é composto por elementos de verso, métrica, rima e pelas formas de repetição de sons, proporcionando tanto a formação imaginária de formas, quanto a possibilidade de ter sensações [Silva, 2014]. Além das estruturas gráficas que um poema pode apresentar, este também pode ser formado por um conjunto de versos, formando estrofe. E um verso, é classificado de acordo com o número de sílabas poéticas, conhecidas como métrica.

A métrica, é identificada diante da escansão dos versos, considerando o seu número de sílabas, distribuição dos acentos e das pausas, conforme regras da sílaba tônica, de encontro vocálico entre palavras, e de licença poética. Para cada número de sílabas existe uma nomenclatura especial, e a coincidência entre últimas sílabas de algumas palavras forma a rima.

A rima é a repetição da combinação de sons semelhantes, presente nas estrofes, mais encontrada no final de versos alternados. Além dos elementos anteriormente descritos, existe o ritmo, gerado através da regularidade de posição de sílabas tônicas e átonas existentes em um conjunto de versos, e Silva [2014] enfatiza que o ritmo é um elemento indispensável em um poema. A identificação do ritmo é feita por meio da separação de sílabas poéticas das palavras de um verso, e assim é possível contar o número destas sílabas e apresentar a medida do verso, para classificação da sua métrica.

2.1.1 Métrica

A unidade de medida de um verso é conhecida por metro, resultado de uma escansão que corresponde à divisão das sílabas poéticas, revelando do texto os seus padrões rítmicos, historicamente consolidados, por meio da posição da sílaba tônica identificada em cada tipo de verso Goldstein [2006]. Com a aplicação da métrica, a rima aparece de forma regular, marcando a semelhança fônica no final de certos versos de determinada estrofe.

A contagem de sílabas métricas obedece aos preceitos de contar as sílabas até a última sílaba tônica, sendo cada verso bem marcado por posicionamento de tônicas em algumas das suas sílabas. Nestas sílabas devem ser considerados alguns fenômenos, como os ditongos crescentes como apenas uma sílaba métrica, e a escolha entre manter ou separar duas ou mais vogais átonas ou tônicas em uma mesma sílaba poética.

Com base na métrica do poema, o verso pode ser classificado como monossílabos, dissílabos, trissílabos, tetrassílabo, pentassílabo, hexassílabo, heptassílabo, Octassílabo, eneassílabo, decassílabo, hendecassílabo, dodecassílabo ou alexandrino, sendo formado de uma a doze sílabas respectivamente. Porém, o verso que possui acima de doze sílabas, é conhecido por verso bárbaro.

A Tabela 2.1 apresenta a metrificação do primeiro verso do poema “Remorso” criado por Olavo Bilac e apresentado na obra de Goldstein [2006]. O verso apresentado, sofreu escansão, ou seja, a separação da sentença em dez sílabas poéticas, com uma alternância entre as sílabas fortes e fracas, e estas são contadas até a última sílaba tônica, e a última sílaba é ignorada. Além disso, foram identificadas duas sílabas fortes nas posições de sexta e décima sílabas, possuindo dois esquemas rítmicos, o primeiro até a sexta sílaba e o segundo até a décima sílaba. Um outro fenômeno ocorrido, é a junção de sílabas com vogais de sons semelhantes, formando assim duas crases.

Tabela 2.1: Sentença Escandida do Poema "Remorso" de Olavo Bilac.

Sin -	to o	que es -	per -	di -	cei -	na -	ju -	ven -	tu (de)
1	2	3	4	5	6	7	8	9	10

A escrita de um poema composto por versos regulares, requer a padronização, mantendo todos os versos com uma mesma estrutura métrica. Por isso, algumas vezes é preciso fazer a junção ou separação de sílabas poéticas em um verso para satisfazer determinada regra clássica estabelecida, fazendo elisão, crase, sinérese, diérese, síncope, conforme explicado a seguir:

- elisão: é a junção de duas sílabas poéticas, devido a ocorrência de uma palavra terminar com vogal átona e a próxima iniciar com vogal ou com a letra h;
- crase: é a fusão de duas vogais com sons iguais;
- sinalefa: é a fusão de duas vogais com sons diferentes;
- sinérese: é a união do hiato em uma só sílaba;
- diérese: é a divisão do ditongo em duas sílabas;
- aférese: é a supressão de vogal no início de uma palavra;
- apócope: é a supressão de vogal no fim de uma palavra;
- síncope: é a supressão de vogal no meio de uma palavra.

A leitura pausada das sílabas poéticas em uma sentença escandida possibilita a identificação de um novo elementos do poema conhecido como ritmo. Este elementos aparece na sílaba que é pronunciada fortemente, e por isso classificada como sílaba tônica, podendo existir mais de uma, formando assim um esquema rítmico.

2.1.2 Ritmo

O ritmo é o elemento essencial do verso. De acordo com Goldstein [2006] o ritmo é formado pela sucessão de unidades rítmicas que resultam da alternância entre sílabas fortes e sílabas fracas, ou entre as sílabas que são constituídas por vogais longas e breves, transmitindo a impressão de musicalidade, e conforme Cunha [2017] a contiguidade de sílabas tônicas desagrada ao ouvido.

Para versos regulares, existe esquema rítmico, e as sílabas acentuadas devem corresponder ao número de sílabas poéticas, para garantir o ritmo no poema, ou seja, para cada número de sílabas de uma determinada métrica, há uma marcação padrão em alternadas sílabas específicas, e de acordo com Goldstein [2006] é exibido abaixo o esquema rítmico mais frequentemente usado tanto para versos regulares, quanto para versos brancos e polimétricos:

- Verso dissílabo - tônica na 2ª sílaba.
- Verso trissílabos - tônica na 3ª; ou na 1ª e 3ª sílabas.
- Verso tetrassílabo - tônica na 1ª e 4ª; ou 2ª e 4ª sílabas.
- Verso pentassílabo - tônica na 2ª e 5ª; ou 3ª e 5ª; ou 1ª, 3ª e 5ª sílabas.
- Verso hexassílabo - tônica na 3ª e 6ª; ou 2ª e 6ª; ou 2ª, 4ª e 6ª; ou 1ª, 4ª e 6ª sílabas.
- Verso heptassílabo - tônica em qualquer sílaba e na última.
- Verso Octassílabo - tônica na 4ª e 8ª; ou 2ª, 6ª e 8ª; ou 3ª, 5ª e 8ª; ou 2ª, 5ª e 8ª sílabas.
- Verso eneassílabo - tônica na 4ª e 9ª; ou 3ª, 6ª e 9ª sílabas.
- Verso decassílabo - tônica na 6ª e 10ª; ou 4ª, 8ª e 10ª sílabas.
- Verso hendecassílabo - tônica na 5ª e 11ª; ou 2ª, 5ª, 8ª e 11ª; ou 2ª, 4ª, 6ª e 11ª sílabas.
- Verso dodecassílabo - tônica na 6ª e 12ª; ou 4ª, 8ª e 12ª; ou 4ª, 6ª, 8ª e 12ª sílabas.

Como exemplo de esquema rítmico, a Tabela 2.2 apresenta um trecho do poema “Remorso”, de Olavo Bilac identificado no livro de Goldstein [2006] com os seus versos são separados em dez sílabas poéticas, sendo classificados como decassílabos. As suas sílabas tônicas são marcadas na 6ª e 10ª posições de cada verso, formando assim a bipartição em dois segmentos rítmicos. Ainda é possível identificar na quadra da Tabela 2.2 uma coincidência de sons nas últimas palavras de cada verso entre as palavras juventude e virtude, e entre as palavras velhice e tolice, caracterizando um tipo de rima cruzada.

Tabela 2.2: Trecho do poema “Remorso” de Olavo Bilac.

<p>Sinto o que desperd<u>icei</u> na juvent<u>ude</u> Choro neste com<u>eço</u> de velh<u>ice</u> Mártir da hipocris<u>ia</u> ou da virt<u>ude</u> Os beijos que não t<u>ive</u> por tolic<u>e</u></p>

Segundo Goldstein [2006] até o início do século XX, os poemas eram compostos com simetria, quando cada poeta exibia o seu estilo através de poemas versificados e influenciados pelo mesmo número de sílabas em todos os versos do poema, considerando a sua acentuação e um determinado tipo de rima. Em meados do século XX, surge uma nova estrutura rítmica em

versos assimétricos, que permite a escrita de poemas de forma solta e distanciada das regras métricas tradicionais, cujo exemplo é apresentado na próxima seção.

2.1.3 Verso

Um verso corresponde à uma linha do poema, e este pode ser classificado como verso regular, verso branco, verso polimétrico, ou verso livre, sendo que cada um apresenta particularidades em sua formação. Em versos regulares, como já apresentado no capítulo anterior, existe uma regularidade na métrica e na marcação de tônicas em todos os versos, além de apresentar rima [Goldstein, 2006].

Os versos classificados como brancos do ponto de vista de Goldstein [2006] obedecem às regras métricas de versificação ou acentuação, favorecendo ritmo aos versos por manter padronizado o posicionamento das tônicas entre estes. A identificação deste verso é facilmente realizada pela ausência de rima, conforme o exemplo exibido na Tabela 2.3, através de uma quintilha do poema “Uruguai”, de Basílio da Gama (séc. XVIII).

Tabela 2.3: Trecho do poema “Uruguai” de Basílio da Gama.

Lá, como é uso do país, roçando Dois lenhos entre si, desperta a chama, Que já se ateia nas ligeiras palhas, E velozmente se propaga. Ao vivo Deixa Cacambo o resto e foge a tempo
--

Apesar da ausência de rima, de acordo com análise feita em Goldstein [2006] trata-se de versos compostos por dez sílabas poéticas, sendo o primeiro, o terceiro e o quarto verso são sáficos, considerados assim devido ao esquema rítmico marcado nas sílabas 4, 8 e 10 de cada verso, os demais são versos heroicos, cujo esquema rítmico é marcado nas sílabas 6 e 10.

Os versos polimétricos sob a ótica de Goldstein [2006] são compostos por versos regulares, de tamanhos diferentes, mas com sílabas fortes, marcadas nas posições indicadas pelas métricas tradicionais, ocupando posições fixas. Estas características podem ser observadas na Tabela 2.4, no poema de Vinícius de Moraes quando cada verso possui um tamanho diferente.

Tabela 2.4: Trecho do poema “Paisagem” de Vinícius de Moraes.

Subi a alta colina Para encontrar a tarde Entre os rios cativos A sombra sepultava o silêncio. Assim entrei no pensamento Da morte minha amiga Ao pé da grande montanha Do outro lado do poente.

Já os versos livres não obedecem a regras estabelecidas para métrica, posição da sílaba forte, nem a regularidade ou mesmo a presença de rima. Eles possuem ritmo irregular, os acentos podem mudar de lugar de acordo com a leitura, a simetria perde espaço para o contraste, a dissonância, o efeito imprevisível ou inesperado. E de acordo com Goldstein [2006] o modo de

compor do artista, traduz a visão de mundo de cada época, como visto no terceto exibido na Tabela 2.5 o poema do século XX de Fernando Pessoa. Como análise de Goldstein [2006] este poema apresenta o “ritmo irregular, dinâmico como o da vida do homem contemporâneo”.

Tabela 2.5: Trecho do poema Fernando Pessoa.

O espelho reflete certo: não erra porque não pensa. Pensar é essencialmente errar. Errar é essencialmente estar cego e surdo.

Como visto nos poemas apresentados, um poema é formado por um conjunto de versos, e a separação entre cada grupo de versos é considerada como estrofe. Esta separação é marcada por uma linha em branco, e o número de versos de cada estrofe corresponde a um determinado tipo de poema regular.

2.1.4 Estrofe

O verso é considerado por Goldstein [2006] como a primeira unidade no interior de um poema, e a estrofe é a segunda unidade. Logo, a estrofe é um conjunto de versos, separados graficamente com espaços em branco, tendo de um a dez versos, nomeados como monóstico, dístico, terceto, quarteto, quinteto, sexteto, sétima, oitava, novena e décima respectivamente. Além disso, ela também é marcada fonicamente pelas rimas. Goldstein [2006] destaca um tempo de predominância no uso de determinados tipos de versos, que variavam de acordo com a época, mas com a chegada dos versos livres, as estrofes são apresentadas com os mais variados tipos, e até mescla de alguns tipos de estrofe.

Em composição popular, a rima do quarteto é restrita a versos pares, o que não ocorre com o sexteto, cujas estrofes podem rimar livremente, bem como a quintilha. A décima é organizada com um quarteto, seguido de um sexteto, obedecendo a um esquema de diferentes rimas em cada uma de suas subestrofes.

2.1.5 Rima

Como abordagem feita em Goldstein [2006] a rima é um parentesco sonoro que pode ocorrer entre dois ou mais versos, seja no final destes, ou no interior dos mesmos, causando impacto musical e rítmico que contribui para dar sentido ao texto. Em um poema, a rima possui classificações, que podem ser identificadas quanto a sua posição no verso, quanto a semelhança de letras, quanto a distribuição ao longo do poema, quanto a sua categoria gramatical e quanto a extensão dos sons que rimam.

A posição da rima no verso ocorre na parte interna ou externa. A rima externa ocorre sempre no final das palavras de cada verso do poema, e por ser explorada neste trabalho, será exemplificada nas próximas seções. Já a rima interna, conhecida como rima coroadada, pode ser identificada com a ocorrência de duas palavras rimando, estando uma no final do primeiro verso e a outra no interior do próximo verso, ou ambas estando no mesmo verso, ou ainda, mesmo estando no mesmo verso, uma aparece imediatamente após a outra. A Tabela 2.6, apresenta uma quadra do poema de Tomás Ribeiro, com um dos tipos de rima interna entre palavras em destaque, combinando o som das palavras: errante e distante, infinda e linda, proporcionando uma sonoridade no poema.

Tabela 2.6: Rima coroada.

Anjo sem pátria, fada errante , Perto ou distante que de mim tu vais, Tá de seguir-te uma saudade infinda , Hebreia linda , que dormindo está.

Na ocorrência de rima interna entre palavras do mesmo verso, pode ocorrer uma pequena pausa, geralmente usada com valor dramático. Este tipo de rima interna é conhecida como rima leonina, exemplificada na Tabela 2.7, através do terceto do poema cujo autor é desconhecido. Nesse caso há uma combinação de sons entre as palavras destacadas: essas e peças, amigo e digo, verso e despeço.

Tabela 2.7: Rima leonina.

Sejam essas , mesmo que na peças A este amigo , as últimas palavras que digo Em livre verso , amigo, me despeço

Uma outra forma de rima interna é “com eco”, muito usada em poesia latina da Idade Média. Este tipo de rima utiliza o recurso de rimar palavras que estejam juntamente posicionadas, conforme exemplo apresentado na Tabela 2.8, no poema de Castro Alves, que apresenta rima entre as palavras: donzela e bela, inspira e lira, canto e santo, bardo e cardo, tremenda e senda, estanca e arranca.

Tabela 2.8: Rima com eco.

Donzela bela , que me inspira a lira Um canto santo de fremente amor Ao bardo o cardo da tremenda senda Estanca, arranca-lhe a terrível dor.

A rima também pode ser consoante e toante, caracterizadas pela semelhança dos sons. Na rima consoante, também conhecida como soante, tal semelhança é representada por consoantes e vogais, ou seja, as palavras que rimam, são repetidas, ou parte delas coincidem em muitas de suas letras, contadas do final para o início de cada uma delas. A Tabela 2.9 apresenta este tipo de ocorrência, com o exemplo da quadra do poema de Carlos Drummond de Andrade, cujas rimas externas, aparecem entre as palavras cadeira e maneira, com base na combinação das palavras “jantar” e “lugar”.

Tabela 2.9: Rima consoante.

Senta-te nesta cadeira e aceita nosso jantar . Tranquilo: em casa mineira nunca faltou um lugar .
--

A rima toante ou assoante, apresenta semelhança apenas entre as vogais tônicas, ou das vogais a partir das tônicas. Desta forma, havendo coincidência entre as consoantes a partir das vogais tônicas, a rima passa a ser considerada perfeita. A Tabela 2.10 apresenta uma quadra popular, com ocorrência de rima entre as letras “i” e “a” das palavras lima e bonita, e entre as letras “o”

e “a” das palavras amora e nora, sendo que nestas últimas palavras, há uma coincidência entre as sílabas inteiras “ora”.

Tabela 2.10: Rima toante.

Tanto limão, tanta lima , Tanta silva, tanta amora , Tanta menina bonita ... Meu pai sem ter uma nora !
--

A rima vista quanto a distribuição ao longo do poema é classificada como cruzada, emparelhada, interpolada e mista. Para isso, existe uma convenção designada à identificação de ocorrência de rima em um poema, relacionando cada letra maiúscula do alfabeto, a uma sequência de versos do poema [Goldstein, 2006]. De acordo com este modo de distribuição ao longo de uma estrofe, estas rimas podem ser cruzadas, emparelhadas, interpoladas e mistas, determinadas pela posição dos versos para cada tipo de rima.

De acordo com as abordagens de Goldstein [2006] a rima cruzada também conhecida por rima alternada, ocorre quando a última sílaba da última palavra do primeiro verso se assemelha a última sílaba da última palavra do terceiro verso, e a última sílaba da última palavra do segundo verso se assemelha a última sílaba da última palavra do quarto verso. Sempre fazendo rima entre versos ímpares e entre versos pares, havendo uma alternância entre rimas, conforme visto na Tabela 2.11, com exemplo da quadra de Carlos Drummond de Andrade, na sequência ABAB, com rimas entre as palavras poeta e pateta, escolar e rimar.

Tabela 2.11: Rima Cruzada.

Versos	Rima
Mas que dizer do poeta	rima A
Numa prova escol ar ?	rima B
Que ele é meio pateta	rima A
E não sabe rimar ?	rima B

Uma outra forma de combinação de rimas é apresentada por Cunha [2017] através do poema de Alberto de Oliveira na Tabela 2.12, composto por rima emparelhada, rimando a cada dois versos de maneira sequencial, conforme relacionamento entre os versos AABBCC, com rimas entre as últimas palavras de cada verso, marcadas pela coincidência de sons entre as palavras recordação e chão, gemia e dia, flor e amor.

Tabela 2.12: Rima Emparelhada.

Versos	Rima
Ele deixava atrás tanta recordaç ão !	rima A
E o pesar, a saudade até no próprio chão ,	rima A
Debaixo dos seus pés, parece que gemia ,	rima B
Levantava-se o sol, vinha rompendo o dia ,	rima B
E o bosque, a selva, o campo, a pradaria em flor	rima C
Vestiam-se de luz, como um peito de amor .	rima C

Na rima interpolada, também conhecida como rima oposta, conforme Cunha [2007] neste tipo de rima ocorre a combinação entre versos internos e outra entre versos externos, seguindo o

relacionamento de ABBA, e pode ser conferido na Tabela 2.13. A rima interna exibida é conhecida como leonina, que ocorre entre as palavras pranto e lento no segundo verso, e a outra rima interna ocorre entre a palavra final do primeiro verso e a palavra deslizando que aparece no meio do segundo verso.

Tabela 2.13: Rima Interpolada.

Versos	Rima
Saudade! Olhar de minha mãe rez <u>ando</u>	rima A
E o pranto lento deslizando em <u>fi</u> o...	rima B
Saudade! Amor da minha terra... O <u>ri</u> o	rima B
Cantigas de águas claras soluç <u>ando</u> .	rima A

Já na rima mista ocorre a combinação entre vários versos, sem que haja uma sequência ordenada, como aborda Cunha [2017]. Pois como exhibe a Tabela 2.14, o 1º verso rima com o 3º, o 2º rima com o 4º e com o 6º, o 5º rima com o 7º e com o 9º, e assim por diante, como visto no poema “Uma criatura” de Machado de Assis. Neste poema aparecem quatro variáveis, sendo que três delas mantêm uma sequência de rima, como ABABCBC, desta forma as palavras que rimam são formidável com insaciável, outra rima de estranhas – que aparece duas vezes - com montanhas, e última rima de abismo, despotismo e egoísmo, e ainda aparece uma palavra que não possui rima, e por isso é chamada rima órfã.

Tabela 2.14: Rima Mista.

Versos	Rima
Sei de uma criatura antiga e formid <u>ável</u>	rima A
Que a si mesma devora os membros e as entr <u>anhas</u>	rima B
Com a sofreguidão da fome insaci <u>ável</u>	rima A
Habita juntamente os vales e as mont <u>anhas</u>	rima B
E no mar, que se rasga, à maneira de ab <u>ismo</u> ,	rima C
Espreguiça-se toda em convulsões estr <u>anhas</u> .	rima B
Traz impresso na frente o obscuro despot <u>ismo</u> .	rima C
Cada olhar que despede, acerbo e mavioso,	
Parece uma expansão de amor e de ego <u>ísmo</u> .	rima C

Como a rima é a coincidência entre sons, algumas vezes esta pode aparecer entre palavras semelhantes como apresentado por Cunha [2017] e exibido na Tabela 2.15, com a coincidência de sons entre as palavras trouxe, doce e fosse, e na palavra sul que rima com azul. Esta estrofe representa uma quintilha de rima perfeita, pois para cada palavra há um som coincidente em forma de ABAAB.

Tabela 2.15: Rima Perfeita.

Versos	Rima
Céu puro que o sol <u>trouxe</u>	rima A
Claro de norte a <u>sul</u> ,	rima B
O teu olhar é <u>doce</u> ,	rima A
Negro assim, qual se <u>fosse</u>	rima A
Inteira <u>mente azul</u> .	rima B

As rimas recebem classificação quanto a extensão dos sons que rimam, representada pelo posicionamento do acento tônico na última palavra de cada verso, e conforme Goldstein [2006] elas podem ser aguda, grave ou esdrúxula, ocorridas respectivamente nas palavras oxítonas, paroxítonas e proparoxítonas como apresentado na Tabela 2.16. O poema intitulado “Poemeto irônico” de Manuel Bandeira apresenta em seu quarteto de rima cruzada, com esquema ABAB, sendo todas as rimas consoantes, com rimas agudas nas palavras paixão e vão, e graves, nas palavras curiosidade e irreabilidade.

Tabela 2.16: Rima por Acentuação.

Versos	Rima
O que tu chamas tua paixão, é tão somente curiosidade E os teus desejos ferventes vão Batendo asas na irreabilidade	Aguda grave aguda grave

Uma rima]também pode ser conceituada como rima rica ou rima pobre, sendo a primeira estabelecida pelo critério fônico, ou seja, as palavras que rimam não possuem a mesma classe gramatical. Já na rima pobre é feita seguindo critério gramatical, e como visto em Cunha [2017] são consideradas rima pobre, as que aparecem em terminações muito usadas no português classificadas como infinitivos, participípios, gerúndios, diminutivos, advérbios, adjetivos, substantivos, e palavras primitivas com seus derivados por prefixação: amor – desamor, ver – rever, entre outros.

Além da rima, o paralelismo de ritmo tanto pode aparecer em versos regulares de um poema, quanto em versos regulares de prosas literárias brasileiras. Porém, a ocorrência de tais paralelismos em prosa costuma ser de difícil identificação, por necessitar de identificação de uma certa regularidade de estrutura métrica, realizada a partir da escansão de suas frases. Dessa forma a próxima seção apresenta o trabalho manual realizado por Almeida [1946] e Campos [1996] na busca de estruturas métricas em versos regulares.

2.2 Estrutura Métrica em Prosa Literária Brasileira

A poesia pode estar presente em obras literárias brasileiras, e esta ocorrência classifica a obra como prosa poética. Conforme Goldstein [2006] esta expressão indica que um trecho é organizado de forma similar a um poema. Esta característica é atribuída com a identificação de estruturas métricas, cujas frases apresentam uma regularidade de esquema rítmico, e as sílabas tônicas são marcadas nas mesmas posições. Tais frases podem aparecer no início, no meio ou em final de parágrafos de uma prosa literária.

Na escola literária modernista, há um destaque para o escritor Euclides da Cunha, que descreve em sua prosa literária as condições de vida do povo nordestino, as características da cidade, as crenças e costumes da população, doenças, miséria e violência geradas no confronto, reprimindo um povo guerreiro que lutava por melhores condições de vida, com base na redução de impostos cobrados.

Campos [1996] relata que Euclides da Cunha presenciou todo o movimento da guerra ocorrida na cidade de Canudos na Bahia, por ter sido correspondente do jornal O Estado de São Paulo. Ele foi designado a escrever um texto jornalístico como cobertura do acontecimento, fazendo uso de figuras de linguagem, repetições e omissões de simbolismos gramaticais. Dessa forma, despertou a curiosidade de outros artistas, que quatro décadas depois de publicado, foram reveladas variadas estruturas de versificação, por poetas em pesquisa ao seu livro.

Ao examinar minuciosamente a prosa de Os Sertões, Campos [1996] declara que “As palavras poesia, poema, empregadas no sentido amplo, emergem instintivamente à leitura do livro, sinalizando o viés estilístico que nos impede de enquadrá-lo *tour court* como prosa”. Além disso, ele exalta o artigo de Almeida [1946] intitulado de “A Poesia d’Os Sertões”, devido a sua descoberta no interior do livro, um número considerável de versos metrificados, sugerindo um novo estilo de texto literário.

Na análise do livro feita de forma manual, os poetas comprovaram a existência de quinhentos decassílabos e dodecassílabos significativos, que são predominantes na literatura poética tradicional brasileira, além de variados gêneros e ritmo, explicitando a presença de poesia. Curiosamente, estes versos estavam posicionados em finais de parágrafos, induzindo o leitor a imaginar que o propósito tenha sido o de valorização das suas ideias em seus terminais, de maneira bem articulada.

A tarefa de busca por estrutura métrica em prosa literária brasileira já pode ser feita de forma automatizada pelo sistema MIVES [Carvalho; Loula E Queiroz, 2016]. Nessa tarefa, o sistema separa segmentos frásicos em sílabas poéticas, classifica-as quanto ao número de sílabas, marca as suas sílabas tônicas e apresenta o seu esquema rítmico. O sistema MIVES portanto, é capaz de identificar automaticamente dentro de uma obra em prosa, as sentenças que possuem estrutura métrica, característica típica dos poemas.

2.3 Fundamentação Computacional

Esta seção apresenta requisitos relevantes para geração de poema computacional, tendo como base, o *survey* elaborado por Oliveira [2017] sobre geração de poesia inteligente, com apresentação de trabalhos realizados por alguns pesquisadores de outros países. Neste sentido, Oliveira [2017] faz uma abordagem sucinta, sobre particularidades de geradores de poemas nos respectivos idiomas, sobre os aspectos relacionados às características mais comuns, sobre às diferentes técnicas aplicadas, e apresenta algumas das avaliações feitas sobre os resultados alcançados. Outros trabalhos também são referenciados nessa seção, com o intuito de melhor explicar cada um destes itens.

2.3.1 Idiomas

Apesar da geração de poema ser teoricamente similar em sua forma, ela sofre diferentes adaptações nas variadas linguagens, devido a articulação dos seus elementos mínimos da fala, pois o som da palavra pode ou não depender da sua forma escrita. Neste sentido, a geração de poemas é guiada por diferentes tradições poéticas, muitas vezes com foco em aspectos diferentes relacionados à forma, conforme o tipo de poema. Mas, de acordo com as características comuns, os *softwares* geradores de poemas desenvolvidos em cada país basicamente são capazes de analisar aspectos fonéticos, léxicos e semânticos, sendo o último

item considerado um ramo da linguística que nem sempre é analisado, muitas vezes pela coerência existente na fonte a ser explorada.

Dos níveis de linguagem presentes no poema, o que melhor caracteriza este texto é o reconhecimento pelo seu estilo. Por isso, Oliveira [2017] reconhece um texto poético pela regularidade da métrica, ou seja, pelo número de sílabas de cada verso, com padrões de sílabas tônicas e pela rima. A modelagem do metro foi considerada nos sistemas de Gervás [2001], de Manurung [2004] e de Chrismartin, Tobing e Manurung [2015] inclusive com indicação da sílaba forte. O fenômeno da repetição de mesmo som também foi explorado pelos mesmos pesquisadores, de modo a identificar semelhança entre as letras ou sílabas das palavras que rimam.

Os aspectos fonéticos, léxicos e semânticos são considerados por Oliveira [2012] como uma tarefa complexa ao desenvolver um gerador automático de poemas para português europeu, e complementa que esses níveis não precisam ser rigorosamente obedecidos. Por outro lado, Oliveira [2012] reconhece que poesia deve conter um alto nível de fenômenos linguísticos interdependentes sobre rima, métrica e ritmo, e foi com base nestes fenômenos que ele desenvolveu um sistema gerador de *haiku*, e de soneto, considerando o número de versos para compor cada estrofe para cada um deles.

Considerando que os sistemas desenvolvidos anteriormente geram poemas em um determinado idioma, Oliveira *et. al.* [2017] desenvolveram um gerador de poemas para os idiomas português europeu, espanhol e inglês. Esse sistema teve como base uma arquitetura modular capaz de definir regras de gramática para cada módulo com um alto nível de personalização. Os parâmetros escolhidos pelo usuário são relacionados a semântica, a estrutura do poema, ao conjunto de palavras consideradas sementes, a estrutura léxica e ao sentimento transmitido ao usuário.

Em seus geradores de poemas Gervás [2000] vem utilizando a fonética do espanhol até os dias atuais, por considerá-lo como simples, devido à fácil identificação das sílabas fortes. Este aspecto possibilita a obtenção de palavras escritas, cumprindo e aplicando as regras do idioma, para gerar romances, quartetos e tercetos. Um dos tipos de poema em espanhol gerado no WASP (*Wishful Automatic Spanish Poet*) é o romance, cujas características são bem definidas para sua forma, por conter versos de cinco ou sete sílabas, e rima com repetição de sons vocálicos.

No WASP são gerados quartetos, e tercetos, e estes seguem as regras de rima do poema em português brasileiro. Gervás [2001] desenvolveu o ASPERA (*Automatic Spanish Poetry Expert and Rewriting Application*), para gerar poesia de forma interativa, e anos depois Gervás [2016] desenvolveu o Hafez para gerar um grande número de poemas, mas com a intenção de realizar uma mudança nos parâmetros estéticos dos resultados obtidos. Neste mesmo ano Ghazvininejad *et. al.* [2016] também desenvolveu um sistema para gerar sonetos a partir de letras de músicas espanholas.

Um dos *softwares* gerador de poema basco foi desenvolvido por Agirrezabal *et. al.* [2013]. O poema deste idioma segue a tradição *bertsolaritza*, semelhante ao repente brasileiro, cujos versos são criados em forma de improviso, seguindo rigorosas limitações de métrica e rima. Uma forma comum de poema é o *zortziko txiki* ou “pequeno de oito”, sendo a primeira linha é formada por oito sílabas poéticas, seguida de sete sílabas, e quatro destas linhas formam uma estrofe.

Na Índia, uma das pesquisas foi realizada por Das e Gamback [2014] que desenvolveram um *software* para geração e poesia em Bengali, língua indo-ariana, considerada pelos autores como

rica em morfologia, sintática e parcialmente fonêmica, isto quer dizer que o seu estilo de pronúncia depende da ortografia e em parte, da fala. A etapa inicial do seu sistema envolve a compreensão da rima de uma linha de entrada fornecida pelo usuário, fazendo assim a identificação das sílabas abertas e fechadas, para compreensão do ritmo e estrutura métrica.

Um dos geradores de poemas originalmente desenvolvido para o idioma inglês, foi o de Manurung [2004] seguido do sistema de Chrismartin, Tobing e Manurung [2015]. Por ser uma língua extremamente fonética, eles fizeram uso de um dicionário de pronúncia para identificar rima em sílabas finais de palavras com o mesmo som. Apesar de Ghazvininejad *et. al.* [2016] também terem desenvolvido para a língua inglesa, eles fizeram adaptação no seu *software* para o espanhol, por ser uma língua menos fonética.

Ghazvininejad *et. al.* [2016] considera a identificação de sílabas fortes mais fácil no idioma espanhol, visto que neste caso muitas palavras são acentuadas, e no idioma inglês estas são identificadas unicamente pela sua pronúncia. Este último sistema citado, conta com um acervo de 97.775 letras de música lírica espanhola. Rashel e Manurung [2014] desenvolveram o sistema Pemuisi para gerar poemas indonésios com base em notícias atuais do idioma indonésio.

Na China, foi desenvolvido o *I, Poet*, por Yan [2016]. Este sistema é capaz de gerar poemas clássicos, com regras e formas bem definidas, contendo uma maior variedade métrica, e exigências de tonicidade e ritmo. Esta modalidade é conhecida como canção, e é um tipo de poesia, que foi explorada também por Zhang e Lapata [2014] analisando grandes repositórios de poemas, para gerar suas variações, fazendo identificação de tônicas e estruturas predefinidas nos poemas gerados.

Apesar da existência de sistemas computacionais na geração de poemas visuais serem datadas desde meados do século passado em tantos outros países, nesta seção são apresentadas as contribuições mais recentes, quando os pesquisadores fazem utilização de regras para poemas regulares com base em textos variados para extração de segmentos frásicos e geração de poemas. Um outro aspecto a ser analisado em geradores de poemas automáticos está relacionado à variada fonte dos dados de entrada nos sistemas.

2.3.2 Dados de Entrada no Sistema

Existe uma variedade de fontes de dados usadas para gerar poemas, e os sistemas extraem linhas ou fragmentos completos de poemas que já tenham sido criados por humanos, sendo feita uma recombinação para geração de novos poemas. Uma destas aplicações foi feita por Agirrezabal [2013] que utilizou fragmentos completos dos poemas gerados por Queneau [1961]. Os sistemas que seguem esta linha, geralmente realizam a identificação de métrica, rima, e semelhança semântica no texto tido como base. Outro tipo de fonte também utilizada é a de textos diversos. Se por um lado, esta modalidade elimina a necessidade de tratamento da sintaxe, dispensando o uso de um modelo gramatical, por outro lado, o assunto perde a sua referência, sendo adulterado, por exibir um novo contexto.

Uma outra forma de alimentar o sistema é através do preenchimento de novas palavras, e nesse caso o modelo deve tratar metáforas e frases consideradas chaves de determinado texto, ou pode ser feito com base na extração automática de sentenças do texto. Neste último caso, os poemas são gerados do texto, mas algumas palavras são substituídas por sinônimos, e dessa forma o texto perde a sua originalidade. Uma proposta semelhante foi apresentada por Gervás [2000]

no sistema WASP, cujos versos são gerados a partir do conjunto de dados iniciais, sendo avaliada a aceitação dos resultados de acordo com o método, que evita repetição de palavras, e um outro método que avalia a métrica, e os aspectos gramaticais que envolvem o vocabulário.

Um bloco do soneto clássico espanhol do século XVI de Garcilaso de La Veja, é usado como entrada no sistema de Gervás [2000] em conjunto com padrões estabelecidos para iniciar o seu processo de geração de poema. Inicialmente, o texto é fragmentado em partes menores, usados para produzir os padrões de referência. Em seu ciclo de geração, o sistema escolhe aleatoriamente uma palavra do vocabulário, acrescenta ao projeto atual, testa se o verso resultante satisfaz as condições estabelecidas, e volta à primeira fase do ciclo, caso contrário, o verso é desconsiderado, e conseqüentemente rejeitado.

Sobre os aspectos gramaticais, Gervás [2000] realizou a combinação da ordem das palavras, com frequência dos adjetivos, considerando comprimento do verso, número de sílabas de cada palavra, posição das tônicas, e rima, aplicadas no vocabulário selecionado, garantindo a qualidade no resultado final. O sistema não distingue variações de gênero da palavra, nem tempos verbais, tampouco flexão numérica da mesma, e semântica.

Para simplificar o processo, Gervás [2000] reconhece o impacto promovido sobre a qualidade dos versos resultantes, e que seus resultados contribuem para discriminar entre aplicação de diferentes estratégias. Já no seu sistema ASPID (*Automatic Spanish Poet Initial Development*), Gervás [2000] implementou cálculos de similaridade entre versos, para estabelecer prioridade alta ou baixa sobre vocabulário completo disponível.

Para desenvolver o ASPERA, Gervás [2001] fez uso das melhores estratégias aplicadas no WASP com objetivo de gerar poesia formal, a partir de uma mensagem em forma de prosa, com especificações sobre o tipo de poema desejado, considerando o número de linhas, forma de rima rígida ou flexível, ambiente urbano ou rural, e o humor, que é interpretado como positivo ou negativo. O sistema então busca nas listas preenchidas com palavras que são separadas por classes gramaticais, e gera o poema, armazenando em um banco de dados, permitindo posteriormente a sua modificação.

Apesar de Gervás [2001] reconhecer que os aspectos de métrica, rima, aliteração e linguagem figurada desempenham papéis importantes, nem todos eles são minuciosamente abordados nos diferentes níveis de linguagem em seu sistema. Um exemplo disso, é que muitas vezes para atender à restrições métricas, é necessário simplificar as regras sintáticas e semânticas. Insatisfeito com os resultados alcançados Gervás *et al.* [2002] fizeram alterações no sistema, sobre contagem silábica, identificação de posição da sílaba tônica, rima e aliteração, possibilitando a combinação destes fenômenos posteriormente. Este sistema recebeu o nome de COLIBRI e seus dados de entrada eram textos em espanhol fornecidos pelo usuário.

Um novo trabalho de Gervás [2016] surge da encomenda de poemas gerados computacionalmente pelo seu sistema WASP, para compor o capítulo de um livro a ser publicado no México. Devido a exigência do tema ser mexicano, então Gervás [2016] utilizou como entrada, um conjunto de poemas de vinte e três autores mexicanos compilados a partir da internet, e combinados a um conjunto de trezentos e quarenta artigos de notícias extraídas das páginas de sites de um jornal mexicano diário.

Um sistema relevante para ser mencionado é o Tra-la-lyrics Oliveira [2005] pois, apesar de gerar letras de músicas a partir de determinada melodia, este lida com identificação de divisão silábica, posicionamento de sílabas tônicas, ritmo, rima e aliteração, ignorando o reconhecimento semântico. Dessa forma, o sistema recebe a melodia, analisa as suas notas,

seleciona as palavras de um vocabulário, que possam ser encaixadas de acordo com o ritmo, por verificar a distância fixa das sílabas tônicas, dando origem a um novo arquivo.

O módulo de vocabulário do Tra-la-lyrics possui três estratégias, com utilização de rima e ritmo, e uso de gramática generativa. Como primeira estratégia, são retornadas palavras aleatórias, na tentativa de obter rima no final de algumas sentenças. Em seguida, dado um conjunto simples de modelos de sentenças do português europeu, é feita a escolha aleatória entre estas, com uma certa coerência sintática e morfológica, analisada pelo sistema. Posteriormente, é feita a verificação das rimas existentes nas sílabas finais das últimas palavras.

Com o intuito de aplicar o domínio semântico, Oliveira [2015] desenvolveu o Tra-la-lyrics 2.0, fazendo uma adaptação do sistema anterior, acrescentando o método de avaliação sobre a correspondência do comprimento da linha, a ocorrência de rimas, e a função de pontuação sobre o número de sílabas átonas em batidas fortes, número de sílabas tônicas em batidas fracas, e sobre palavras interrompidas por pausas na melodia. Cada recurso possui um peso, relacionado a uma sanção, sendo que na ocorrência de rima é aplicado um bônus.

Um outro sistema desenvolvido por Oliveira [2009] foi o PoeTryMe, considerado pelo pesquisador como uma plataforma inteligente, capaz de gerar poesia de maneira automática. Este sistema tem como base de dados as estruturas léxico-semânticas, extraídas de três dicionários da Língua Portuguesa europeia, que servem de referência para avaliação do uso das sentenças, por meio da Informação Mútua Pontual (PMI), que é um cálculo aplicado entre pares de termos de um texto dado como entrada.

Novas funcionalidades do PoeTryMe foram implementadas por Oliveira [2012] cujos dados de entrada são da rede social *Twitter*. O sistema PoeTryMe foi descrito com foco na arquitetura modular por Oliveira *et. al.* [2017] com as devidas adaptações para extensão de multilinguagens, considerando os recursos linguísticos disponíveis para cada um dos idiomas de português europeu, espanhol e inglês.

Dessa forma, os parâmetros estabelecidos para poesia, são referentes aos modelos de linha, tamanho, sentimento, e palavras-chaves são preenchidos pelo usuário, e diante da combinação destes componentes, o sistema gera o poema em português europeu. Além da parametrização apresentada, o sistema PoeTryMe envolve os fenômenos linguísticos de rima, métrica, aliteração e linguagem figurada, que são as técnicas aplicadas, para gerar poesia, sem noção explícita de ritmo, e com sentenças semânticas coerentes, organizadas de acordo com as normas predefinidas.

Outro gerador de poemas que também faz utilização de palavras-chaves é o de Yan [2016] que as recebe como entrada, em forma de tópicos, temas ou cenários, e gera quadras de quatro ou oito linhas, reguladas por palavras de cinco a sete caracteres. Tal abordagem também foi explorada por Zhang e Lapata [2014] operando em um contexto interativo, cujas palavras são relacionadas às frases que devem compor o poema.

Já a fonte de inspiração de Misztal e Indurkha [2014] é texto de conteúdo emocional publicado, oferecendo a escolha de humor conforme conteúdo escolhido, sendo expresso em forma de poemas, fazendo utilização de recursos léxicos, como sinônimos de dicionários. As palavras são selecionadas e organizadas em forma de frases, fazendo uso de uma gramática própria livre de contexto, cuja característica principal é ter todas as regras atribuídas às suas palavras. Porém, uma sentença pode ser escrita de forma correta, mas não ter significado.

O corpus utilizado por Das e Gamback [2014] foi extraído de vários poemas Bengali publicados no *site Kobita*, priorizando os poemas que foram escritos para criança para facilitar o seu

trabalho. Desta forma, foi feita a divisão das palavras em sílabas poéticas pra identificar as unidades rítmicas, e a fragmentação em letras para identificação de vogais e consoantes. Assim, para geração das rimas, o ritmo é preservado na escolha das sentenças.

Para geração de poesia pentâmetro iâmbico, considerado como tipo de métrica usada para descrever o ritmo do poema e cinco sílabas poéticas, Ghazvininejad *et. al.* [2017] usaram 125.074 sinônimos do dicionário de CMU, fazendo identificação e atribuição de valores 0 ou 1 para sílabas átonas e tônicas respectivamente, fazendo a extração apenas daquelas que atendem ao padrão iâmbico de ritmo, totalizando 106.019 palavras, desconsiderando as palavras ambíguas monossilábicas.

As ferramentas apresentadas até o momento foram implementadas com a utilização de técnicas que atendessem a cada necessidade, conforme a exigência feita de acordo com a modularização do sistema, compatíveis com as regras estabelecidas para o seu pleno funcionamento. Logo, a próxima seção aborda as escolhas e estratégias de programação usadas por cada pesquisador.

2.3.3 Técnicas Aplicadas

Algumas das técnicas computacionais utilizadas para desenvolver *softwares* geradores de poemas são do domínio da Inteligência Artificial (IA) pela ideia de tornar este processo mais criativo e aperfeiçoado, com uso da linguagem natural extraído de gramática. As técnicas são: Raciocínio Baseado em Casos (RBC), Tabela de geração, Modelo Estocástico *Hill Climbing*, Algoritmo Evolutivo e variados tipos de Redes Neurais. Alguns destes sistemas são capazes de criar os seus próprios versos, a partir de palavras informadas em uma interface, e outros deles, são gerados a partir de textos, de acordo com as especificações feitas por usuários, com base em regras literárias e gramaticais, para obter aprovação do leitor, pelos versos gerados.

2.3.3.1 Raciocínio Baseado em Casos

A técnica de Raciocínio Baseado em Casos (RBC) permite a recuperação, reutilização, revisão e retenção de uma experiência passada em um domínio de problema, para tratar situações novas, nas mais variadas áreas do conhecimento. Esta técnica foi explorada por Gervas [2000], em seu sistema WASP (*Wishful Automatic Spanish Poet*), tendo como dados de entrada um conjunto de palavras e um conjunto de padrões de versos, retornando um poema. O seu algoritmo segue o padrão de gerar e testar, para garantir que as condições métricas estão sendo atendidas, e em caso contrário, o verso considerado rascunho é descartado e um rascunho para o novo verso é iniciado.

No âmbito da geração de poemas, Gervás [2001] aplicou esta técnica também no ASPERA (*Automatic Spanish Poetry Expert and Rewriting Application*), escrito na linguagem criada pela NASA, denominada Clipes. Este sistema foi considerado como uma versão melhorada das estratégias de construção desenvolvida no WASP, por possuir soluções inteligentes com chance de superar as técnicas computacionais combinando geração de linguagem natural com técnicas de RBC.

Posteriormente foi também utilizada por Gervás *et. al.*, [2002] aplicada no sistema COLIBRI (*Cases and Ontology Libraries Integration for Building Reasoning Infrastructures*). Em suas etapas, no processo de funcionamento, o usuário informa especificações para o poema, e o sistema pesquisa uma estrofe que atenda a estas especificações, com base no seu *corpus* com exemplos de versos. O sistema funciona recuperando o vocabulário e linhas consideradas como exemplos, fazendo a reutilização destas linhas, para combiná-las às palavras do vocabulário, e

apresentar o resultado ao usuário para revisão. Na sequência, o sistema armazena a sentença que considera adequada às futuras produções.

A Figura 2.1 foi adaptada do artigo de Gervás [2001] apresenta com detalhes o uso do RBC aplicado no sistema ASPERA, sendo que cada linha sequencial do poema passa por processos diferentes de forma simultânea, e a fase de revisão e análise garante que as palavras não serão repetidas. No processo de recuperar etapa, devem ser selecionadas palavras e linhas adequadas para cada fragmento do poema, sendo que os exemplos de linhas e padrões já foram atribuídos na fase de planejamento.

Na etapa de reutilização, conforme Gervás [2001] o ASPERA estabelece prioridades a todas as palavras disponíveis conforme critérios designados para coincidência de últimas sílabas, com prioridade alta, e as palavras encontradas no fragmento da mensagem pretendida apresentam prioridade baixa. Tais prioridades podem ser modificadas na geração das linhas iniciais. A próxima etapa é a de reutilização, com aplicação de regras de comprimento da linha, para cada conjunto de fragmento da mensagem e de palavras selecionadas.

Os elementos são formados pelo padrão de linha que é seguido e pelo rascunho da linha atual, sendo que em cada passo da geração as palavras que devem combinar com a próxima linha são escolhidas no padrão POS *tag*, conforme a ordem de prioridade. Caso a palavra escolhida atenda ao requisito, esta é anexada ao rascunho da linha atual e ocorre a iteração do processo, caso contrário, o sistema permite um retrocesso à última palavra e é feita uma nova tentativa.

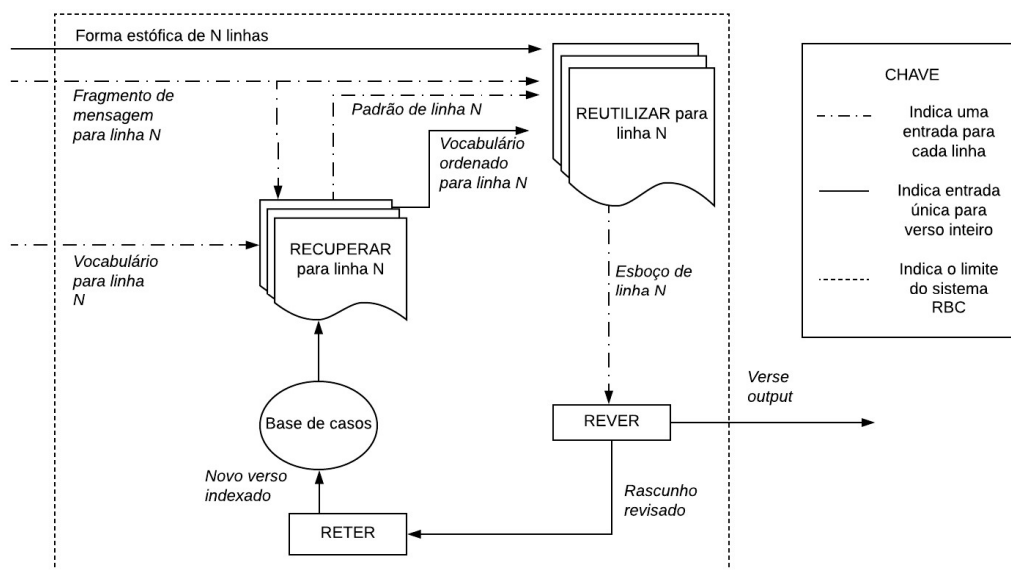


Figura 2.1: Geração de versos usando RBC. Adaptada de Gervás [2001].

Ainda na etapa de reutilização, de acordo com Gervás [2001] o ASPERA pode ser revisado, caso tenha feito o retrocesso e não tenha identificado uma palavra para ser atribuída. A cada modificação de uma determinada palavra, esta passa por nova revisão, e fica restrita às palavras atribuídas a todos os fragmentos anteriores ou posteriores que estejam neste processo.

A etapa de revisão inicia com o término dos processos de geração, e todo o rascunho do poema é apresentado ao usuário, que é obrigado a validar cada uma das linhas, e as suas alterações são aceitas pelo sistema, testando apenas a correção da métrica. Na etapa de manter, é criado um arquivo de dados contendo os exemplos de linhas e padrões correspondentes em formato

correto, juntamente às palavras introduzidas pelo usuário, garantindo a melhora na qualidade dos resultados em função do uso regular do sistema.

Dessa forma, Gervás [2001] implementou o ASPERA para ser capaz de identificar as características mais significativas de casos apresentados, para mostrar uma melhor alternativa como resposta, fazendo o armazenamento do caso e a sua respectiva solução. O alto número de casos relevantes, garante a qualidade do sistema, justificando o uso dessa aplicação, mas quem valida o poema é o usuário. Em seguida, o sistema faz a análise linguística dos poemas corrigidos e adiciona informações do usuário nos seus repositórios para uso subsequente.

Em razão do pedido feito, Gervás [2016] realizou algumas alterações no sistema WASP, pois foi feita a solicitação explícita de poemas genuínos, ou que nunca tivessem sido publicados em outro lugar, para evitar problemas com direitos autorais. Dessa forma, Gervás [2016] inclui um novo procedimento que baseia-se na geração de texto de n-Gram, com um modelo linguístico treinado sobre um *corpus* especificamente construído para corresponder ao objetivo de melhorar as funções da aptidão geral para que apenas os resultados de melhor qualidade fossem aceitas na fase de avaliação.

O outro objetivo foi o de melhorar o procedimento de construção dos poemas gerados. Na versão anterior do sistema WASP as palavras para compor o poema eram escolhidas aleatoriamente, na nova versão tais palavras são escolhidas com base no tema, combinação com o final da frase, controle de repetição das frases já geradas, e restrição ao comprimento das frases.

2.3.3.2 Tabela de Geração

A técnica utilizada por Chrismartin, Tobing e Manurung [2015] é a de tabela de geração, por ser uma abordagem de programação dinâmica para geração de texto. Tal técnica estrutura os dados de entrada em forma de árvore, e armazena tudo o que é gerado. O processamento ocorre do último nó até a raiz, iniciando com a adição de entrada de cada palavra, indexadas de acordo com a ordem que aparecem na frase, fazendo com que este índice acelere o processo de combinação das palavras, uma vez que as arestas incidentes podem ser combinadas para produzir novas arestas, que representam estruturas maiores.

Em seu sistema Chrismartin, Tobing e Manurung [2015] implementou pela primeira vez uma extensão da tabela de geração, levando em consideração a restrição rítmica do poema. Para isso, os pesquisadores usaram o padrão gramatical, considerando uma sílaba fraca, e uma forte, servindo assim como representação às restrições métricas que devem servir como entrada, conforme estrutura apresentada na Tabela 2.17. Caso o padrão não seja atendido, as palavras não são adicionadas na estrutura.

O ponto de partida considerado por Chrismartin, Tobing e Manurung [2015] para construção de padrões de sílabas tônicas, é o da classe léxica, que é identificada a partir do dicionário de pronúncia, que contém as palavras organizadas em lista, classificadas quanto a aspectos morfológicos, semântico, e operador que pode ser léxico ou de predição. O aspecto de predição corresponde à adição de nova palavra que possa combinar com a frase, de maneira a fazer sentido. Similar a esta abordagem, Oliveira [2012] também fez uso da técnica de Tabela de Geração para analisar sentenças de acordo com as entradas gramaticais de arquivos de texto editáveis, fazendo mapeamento direto entre as palavras da relação da tabela e as regras gramaticais.

Tabela 2.17: Exemplos de Entrada na Tabela "the dog saw the cat" [Chrismartin, Tobing e Manurung, 2015].

No.	Phrase	Category	Semantics	Operator
1	<i>Dog</i>	<i>noun(d)</i>	<i>d:dog(d)</i>	<i>Lexical</i>
2	<i>the</i>	<i>det(d)</i>	<i>d:definite(d)</i>	<i>Lexical</i>
3	<i>the</i>	<i>np(d) → det(d) • noun(d)</i>	<i>d:definite(d)</i>	<i>Prediction(2)</i>
4	<i>saw</i>	<i>verb(s, d, c)</i>	<i>s:see(s), arg1(s, d), arg2(s, c)</i>	<i>Lexical</i>
5	<i>saw</i>	<i>vp(s, d) → verb(s, d, c) • np(c)</i>	<i>s:see(s), arg1(s, d), arg2(s, c)</i>	<i>Prediction(4)</i>
6	<i>cat</i>	<i>noun(c)</i>	<i>c:cat(c)</i>	<i>Lexical</i>
7	<i>the</i>	<i>det(c)</i>	<i>c:definite(c)</i>	<i>Lexical</i>
8	<i>the</i>	<i>np(c) → det(c) • noun(c)</i>	<i>c:definite(c)</i>	<i>Prediction(7)</i>
9	<i>the dog</i>	<i>np(d) → det(d) noun(d) •</i>	<i>d:definite(d), dog(d)</i>	<i>(1)+(3)</i>
10	<i>the dog</i>	<i>s(-) → np(d) • vp(-, d)</i>	<i>d: definite(d), dog(d)</i>	<i>Prediction(9)</i>
11	<i>the cat</i>	<i>np(c) → det(c) noun(c) •</i>	<i>c: definite(c), cat(c)</i>	<i>(6)+(8)</i>
12	<i>the cat</i>	<i>s(-) → np(c) • vp(-, c)</i>	<i>c: definite(c), cat(c)</i>	<i>Prediction(11)</i>
13	<i>saw the cat</i>	<i>vp(s, d) → verb(s, d, c) np(c) •</i>	<i>s:see(s), arg1(s, d), arg2(s, c), definite(c), cat(c)</i>	<i>(5)+(11)</i>
14	<i>the dog saw the cat</i>	<i>s(s) → np(d) vp(s, d) •</i>	<i>s:see(s), arg1(s, d), arg2(s, c), definite(c), cat(c), definite(d), dog(d)</i>	<i>(10)+(13)</i>

Como recurso de implementação Oliveira [2012] utilizou o algoritmo de tabela de geração de Earley para analisar sentenças de acordo com as gramáticas, e todas as linguagens livres de contexto extraídas da base de conhecimento léxico de nome Cartão, que contém dados de três dicionários do português espanhol capaz de realizar tríplexes semânticas, fazendo relações entre palavras de mesmo sentido. Um outro recurso usado por Oliveira [2012] foi o SílabasPT3 considerada uma API que realiza divisão silábica e faz a identificação da sílaba tônica da palavra, e contou também com a ferramenta LABEL-LEX4 que fornece informações morfológicas e de parte da fala.

A arquitetura modular do PoeTryMe é uma plataforma de geração de poesia que permite a melhoria independente de cada módulo [Oliveira, 2012]. Tal arquitetura pretende ser versátil fornecendo um alto nível de personalização de acordo com as necessidades do sistema e das ideias do usuário. Assim, é possível definir a semântica a ser usada, os modelos de frase na geração de gramática, a estratégia de geração e a configuração do poema como mostrado na Figura 2.2.

Na segunda versão do PoeTryMe Oliveira [2012] realizou as adaptações necessárias para gerar poemas a partir de textos do *Twitter*, reutilizando a produção de novos segmentos de texto, para destacar a tendência social da rede, interpretando textos produzidos por usuários do *Twitter* ou baseado em relações semânticas extraídas ou deduzidas. Os novos recursos do sistema permitiram a reutilização de fragmentos de *tweets* produzidos pelo usuário, fazendo substituição com palavras sinônimas, além de acrescentar fragmentos que possam destacar a tendência social da rede [Oliveira, 2012]. Os poemas produzidos podem incluir novos fragmentos ou produzidos pelo sistema baseado nas palavras-chaves, mantendo o medidor das frases e fazendo as combinações de rimas.

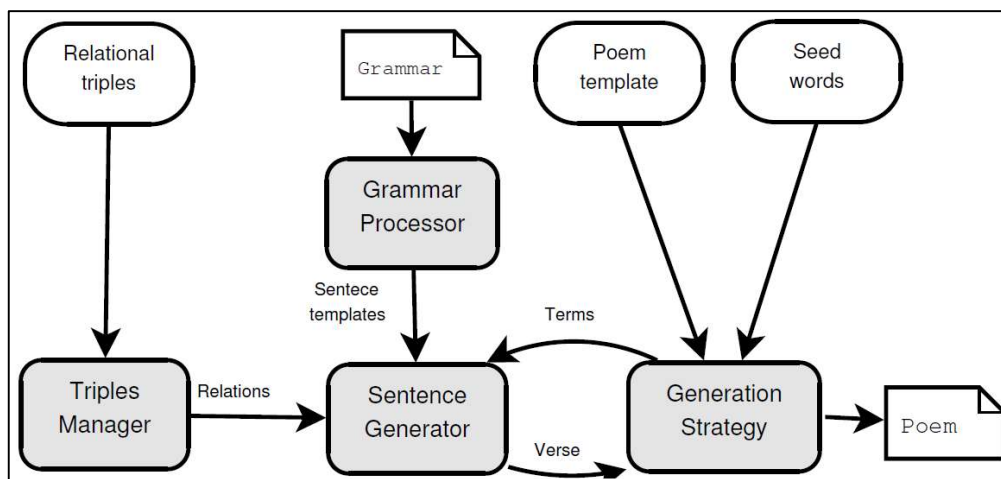


Figura 2.2: Arquitetura do PoeTryMe [Oliveira, 2012].

2.3.3.3 Modelo Estocástico *Hill Climbing*

O modelo estocástico de *hill climbing* foi aplicado por Manurung, Ritchie e Thompson [2000] em seu sistema, que fizeram a escolha pela eficiência, e especialmente por ser um algoritmo considerado como criativo por receber uma coleção de informações tidas como um problema, e realiza a busca pela melhor solução. Por isso Manurung, Ritchie e Thompson [2000] consideram este modelo como um elemento perfeitamente adequado no processo de geração de poemas, sendo que na pesquisa em escalada há uma evolução algorítmica, feita a iteração aplicada a um conjunto ordenado (população) de soluções candidatas (indivíduos) em duas fases: na avaliação e na evolução.

O algoritmo de Manurung, Ritchie e Thompson [2000] inicialmente realiza a busca pela melhor solução de forma randômica, fazendo modificações ao longo das suas iterações, efetuando pequenas adequações no texto. O término do seu processamento se dá com a impossibilidade de encontrar uma melhoria possível em uma iteração, conforme visto na Figura 2.3. Apesar de oferecer uma solução adequada, não existe uma garantia que esta seja a solução ideal.

O modelo utilizado por Manurung, Ritchie e Thompson [2000] é considerado como uma estratégia gulosa, pois faz a busca por uma solução adequada para o momento, seguindo na direção de valores crescentes ou decrescentes da função objetivo. Na implementação em questão, a busca é feita por um possível texto com toda a sua representação subjacente, podendo ocorrer movimento de acordo com o nível de representação de fonética e de semântica.

Manurung, Ritchie e Thompson [2000] descrevem em seu artigo, as dificuldades da geração da poesia, decorrentes dos recursos naturais informativos tradicionais relacionados a NLG (Geração de Linguagem Natural). Eles apresentam a implementação preliminar do NLG, e destacam como deficiências das implementações anteriores de outros pesquisadores na geração de poesia, a falta de sentido apresentada nas obras, e ausência da forma poética relacionada a ritmo, rima e linguagem figurada.

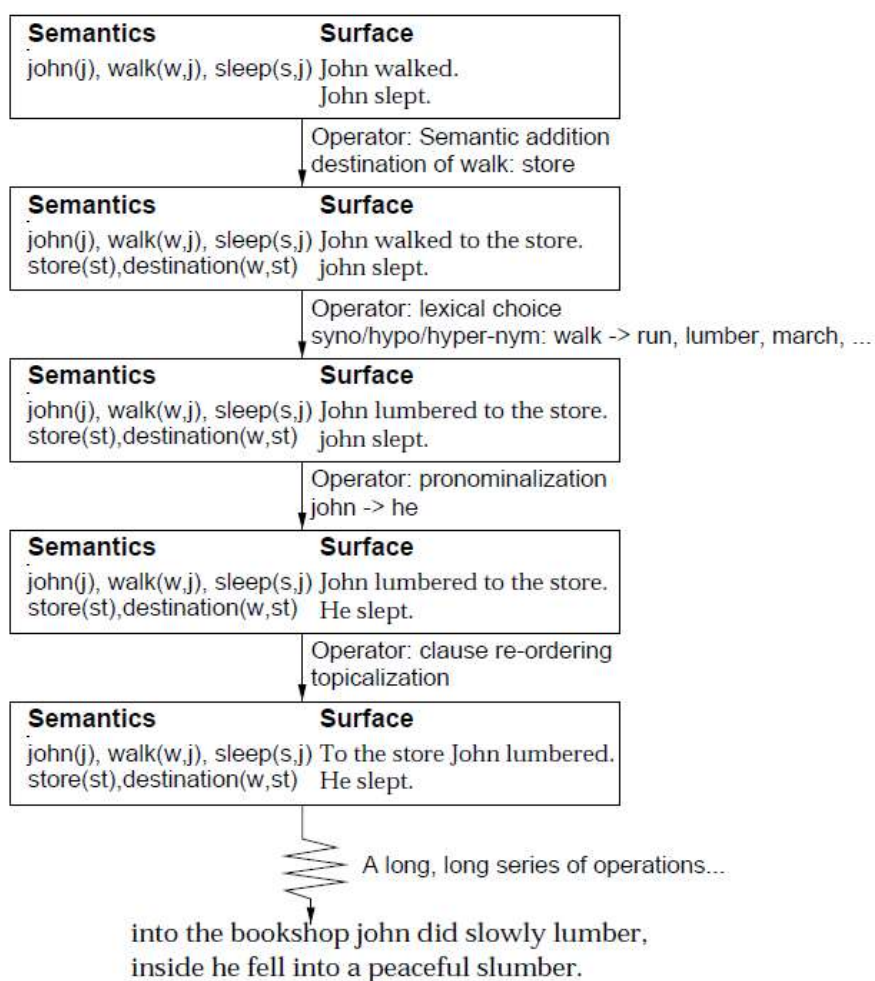


Figura 2.3: Diagrama idealizado de uma pesquisa estocástica [Manurung, Ritchie e Thompson, 2000].

De acordo com Manurung, Ritchie e Thompson [2000] para produzir poemas coerentes, o sistema precisa ter riqueza de recursos como uma ampla gramática que permita parafrasear, fazer tratamento léxico que forneça informações fonéticas e uma base de conhecimento sobre todos estes elementos. As principais características que Manurung, Ritchie e Thompson [2000] buscam sobre o sistema de geração de poesia, são a ocorrência regular de padrões sintáticos e fonéticos como medidor, rima e aliteração, adotando assim uma visão clássica da poesia. Além disso, são oferecidos tratamento semântico e de construção do poema.

Em sua pesquisa Manurung [2004] defende a sua tese sobre uma abordagem evolutiva para gerar poesia através de um método de busca heurística no desenvolvimento do McGonagall. Foi usada representação linguística com base em operações sintáticas LTAG (*Lexicalized Tree Adjoining Grammar*) apropriados para NLG (*Natural Language Generation*) baseada em Algoritmo Evolutivo. Dessa forma, a produção é feita com base na comunicação, na tentativa de alcançar como objetivo o texto, usando as fontes disponíveis de conhecimento linguístico.

O diagrama da Figura 2.4 mostra a arquitetura de *pipeline* para estabelecer a comunicação entre os processos de determinar o conteúdo lido com a semântica proposicional da mensagem, selecionar o conjunto de proposições, planejamento das sentenças, seleção e organização das estruturas sintáticas, para transmissão das proposições selecionadas. Em cada um dos processos

são tomadas algumas decisões como escolha de expressões tidas como referência, e das combinações. Por serem escolhas interdependentes, uma escolha pode ser impedida em detrimento de uma outra escolha, gerando um poema razoável ou até deixando de gerar o mesmo.

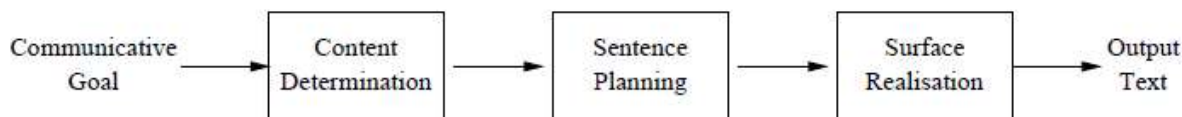


Figura 2.4: Os três processos Reiter em uma arquitetura de pipeline [Manurung, 2004].

A técnica de NLG foi também utilizada por Rashel e Manurung [2014] no desenvolvimento do seu sistema Pemuisi, fazendo uso de um modelo com frases prontas. Como primeiro teste, as frases foram extraídas de um *corpus* contendo 213 poemas de escritores indonésios, e em seguida, foram extraídas frases de artigos de notícias neste idioma. Dessa forma, são feitas combinações de textos de um conjunto de palavras considerado como modelo, e as mesmas são indexadas em uma *tag* léxica (LTAG), de acordo com as classes gramaticais de substantivo, verbo, adjetivo ou pronome.

A área explorada por Agirrezabal et. al. [2013] na implementação do sistema BertsoBOT, foi PNL (Processamento de Linguagem Natural), aliada à técnicas de robótica, com síntese e reconhecimento de fala, para gerar estrofes em basco, automaticamente. Dessa forma, foi feito um analisador morfológico para extração de padrões de estrofes para gerar novas estrofes, seguindo o padrão mais comumente usado.

2.3.3.4 Algoritmo Evolutivo

A proposta do uso de algoritmo evolutivo para desenvolvimento do seu sistema foi feita por Manurung [2004] que reconhece a geração de poesia como uma tarefa incremental, e determinada ideia passa por várias iterações para aperfeiçoamento, gerando uma diminuição da diversidade de versos, ficando apenas aqueles considerados como os melhores qualificados.

A aplicação dessa técnica se dá por meio de estratégias evolutivas, que lidam com problemas de otimização de parâmetros, até que seja identificado o melhor poema. Assim, é feita uma série de iterações, gerando rascunhos descartáveis até que seja identificado o modelo considerado ideal. A ideia é gerar vários poemas iniciais, com base em um método, evoluindo até a identificação do poema adequado, avaliado por pessoas, que tem como padrão o conjunto de recursos de paralelismos relevantes para poesia.

No sistema desenvolvido por Manurung [2004] as mudanças na população são obtidas com base na aplicação de operadores de *crossover* e mutação, sendo que o primeiro operador gera novos poemas a partir de outros dois poemas na população, e Manurung [2004] considera a sintaxe do primeiro poema e as palavras ou rima do último, e realiza a permuta entre partes do primeiro poema e partes do último.

Quanto a mutação, esta pode envolver a substituição de palavras por todo o poema, em apenas uma determinada linha, ou mudando o tipo de rima Manurung [2004]. Considerando que na mutação também é possível adicionar, excluir ou alterar conteúdos no poema, Manurung [2004] fez a opção por valorizar a semântica.

2.3.3.5 Quadro Negro

O quadro negro é uma técnica análoga a um espaço de trabalho comum, compartilhado por especialistas para apresentar soluções parciais e outras informações sobre o problema em vários níveis de abstração conforme Misztal e Indurkha [2014]. Esta técnica foi usada por Misztal e Indurkha [2014] para gerar poemas que recebe os dados de entrada e suas restrições relacionadas a linha, sílabas, padrões de rima e requisitos gramaticais relativos a tempo e pessoa. A Figura 2.5 mostra a arquitetura do quadro negro usada no sistema.

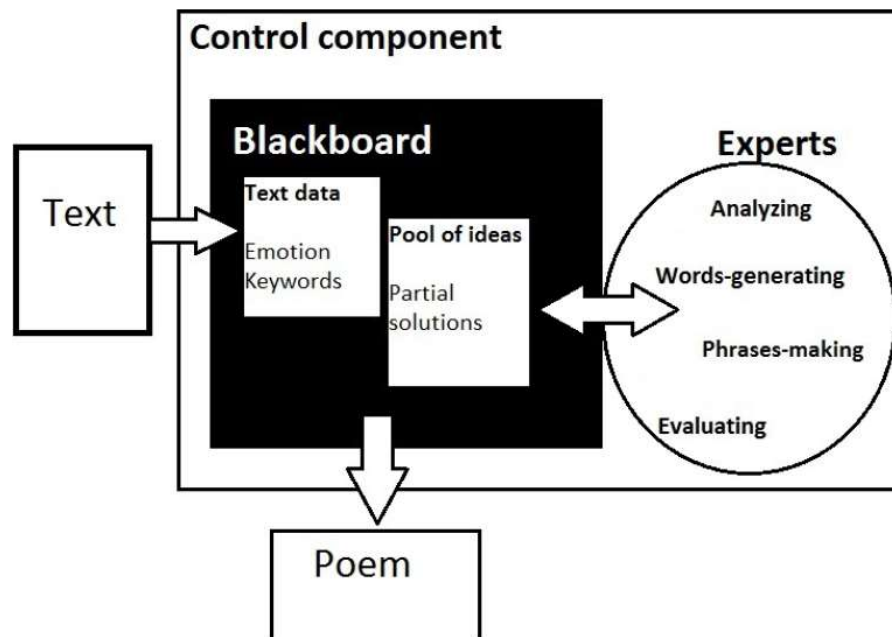


Figura 2.5: Arquitetura do quadro negro usada no sistema de Misztal e Indurkha [2014].

A Figura 2.5 mostra a entrada manual de um texto no sistema, bem como as restrições do seu estilo, ou seja, a parametrização do sistema. Em seguida os dados são analisados, e as informações extraídas do texto são as frases-chave classificadas pelo valor referente ao número de associações que podem ser derivadas por especialistas. O tema do poema é selecionado pela frase considerada mais inspiradora que corresponde ao estado emocional definido como sentimento para o texto.

As expressões derivadas da frase são ampliadas com o uso de conhecimento dos especialistas referente aos recursos lexicais. As palavras vão para uma área conhecida como piscina, e lá elas são classificadas de acordo com a forma gramatical e com o seu significado. Estas palavras podem ser usadas por outros especialistas para produção de novas frases, considerando a parametrização do sistema para geração das linhas do poema.

2.3.3.6 Rede Neural

Uma RNA (Rede Neural Artificial) é conhecida como uma das técnicas computacionais, composta por unidades de processamentos capazes de reproduzir o comportamento humano das suas diversas habilidades cognitivas. De acordo com Rojas [1996], este tipo de rede é construída por meio uma modelagem matemática, composta por canais de comunicação, que são

associados a um determinado peso, que se conectam para projetar modelos abstratos de redes neurais artificiais.

São inúmeras as aplicações que podem ser feitas com base nesta técnica, mas todas elas têm como base a comunicação, combinando propriedades de entrada, produzindo soluções. Utilizada na geração de poemas, o sistema recebe predicados da sua estrutura, bem como o vocabulário a ser usado, sendo necessário restringir rigorosamente o número de linhas, e sílabas por linha, e de forma flexível a escolha da rima.

Como exemplo de aplicação de RNA na geração de poemas, pode-se destacar o modelo de rede neural escolhido por Zhang e Lapata [2014], baseado em *Deep Neural Network* (DNN) foi escolhida por Yan [2016], com inclusão da *Recurrent Neural Network* (RNN). Dadas as características de uma frase, é possível gerar várias linhas de um poema de forma incremental, com base na fonética, a estrutura e a sua semântica. Considerando a possibilidade de haver uma rede neural para selecionar a estrutura da linha e outra para guiar a geração de palavras únicas dentro de uma linha, Yan [2016] vai além, e acrescenta iterações de refinamento de poemas a um processo prévio.

O RNN foi escolhido por Ghazvininejad et al. [2017] para desenvolver o seu programa Hefaz capaz de gerar um grande número de poemas distintos de acordo com o tema fornecido pelo usuário, obedecendo ritmo e restrições de rima guiado por um estado finito para mediar todos os caminhos possíveis das palavras do vocabulário que obedecem às restrições de ritmo formal e palavras que rimam. Para isso, as palavras de entrada são organizadas em uma lista de palavras, com frases relacionadas, pontuadas de acordo com a semelhança entre outras, tendo como saída por exemplo: *coronel (1.0)*, *tenente-coronel (0.77)*, *comandante (0,67)*. Em um outro método, Ghazvininejad et al. [2017] usa informações mútuas pontuais (PMI), fazendo relacionamento de palavra à uma frase. Esta entra na relação de candidatas a compor o poema. No seu terceiro método, as palavras passam por um treinamento, para que sejam selecionadas.

A Figura 2.6 mostra uma visão geral do sistema que define as tarefas de selecionar um amplo vocabulário de palavras, verifica as sílabas tônicas de cada uma delas, e de acordo com o tópico informado pelo usuário, é carregado um conjunto de palavras relacionadas, sendo selecionados os pares que rimam em linhas finais. Em seguida é construindo um caminho com todas as sequências possíveis das palavras do vocabulário que obedecem ao ritmo formal, a seleção do caminho é feita por uma RNN.

[Misztal; Indurkha, 2014] adotaram a multiabordagem com um conjunto de especialistas artificiais, focados no aspecto de geração de poesia, interagem ao compartilhar resultados em um quadro-negro. Estes especialistas contribuem com palavras que correspondem a um determinado tópico ou emoção, organizando-as e selecionando aquelas que apresentam as melhores soluções conforme as restrições estabelecidas.

A geração de poesia também foi abordada por Yan *et. al.* [2013] como uma geração estrutural de poemas formados por quadras, com padrões de tônicas e rimas, número limitado de caracteres por linha, cujas características são utilizadas como forma de otimização, e os poemas candidatos são recuperados por um conjunto de palavras-chaves, que são segmentadas conforme os termos, e agrupados dada a sua semântica. Dessa forma, o sistema utiliza o quadro de sumarização gerativo, fazendo as iterações entre as características poéticas de um dataset com 61.960 poemas chineses.

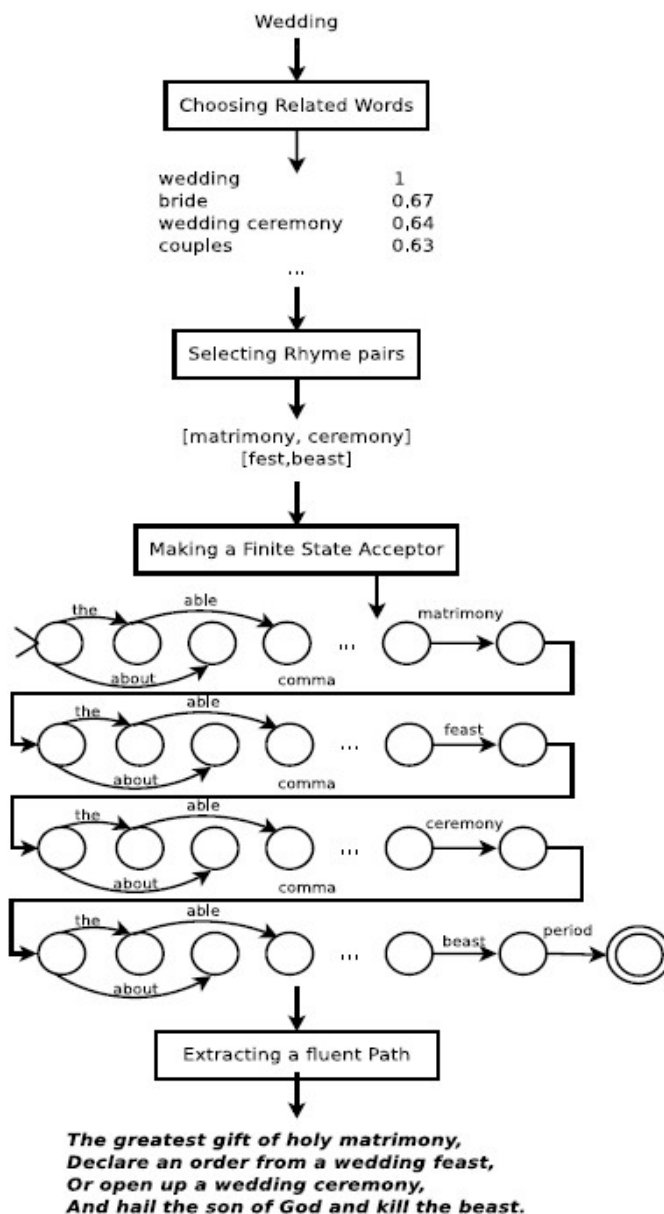


Figura 2.6: Visão geral do Hefaz [Ghazvininejad *et. al.*, 2017].

2.3.4 Avaliações dos Geradores de Poema

Os sistemas desenvolvidos para gerar poemas foram avaliados de acordo com os requisitos considerados relevantes por cada pesquisador. De modo geral, eles consideram os aspectos fonéticos, linguísticos e semânticos. Assim, nesta seção são apresentadas as avaliações e resultados alcançados com as produções computacionais.

A avaliação do sistema WASP ocorreu em duas etapas Gervás [2000]. Na primeira etapa, foram atribuídos três valores a cada poema resultante: número de versos do poema, valor para sua qualificação sintática, valor por sua classificação estética, sendo estes valores usados no primeiro estágio de filtragem com a intenção de otimizar qualidade de versos que não fazem sentido. Na segunda etapa, foi feita uma correção sintática sobre a coerência poética, avaliada

por uma equipe de voluntários que receberam uma lista de 45 poemas para selecionar os cinco melhores, atribuindo uma ordem de preferência, sendo que o maior recebia 5 pontos.

Na última versão do sistema WASP Gervás [2016] realizou uma avaliação formativa sobre o protótipo existente, configurado de forma que uma grande população de rascunho foi construída, com uma série de formas poéticas, avaliadas pela combinação entre a função de aptidão para o tema, as formas poéticas consideradas, e a originalidade do poema. Para isso os módulos de revisão foram desligados para evitar interferência nesse teste.

O esquema de avaliação foi considerado por Manurung, Ritchie e Thompson [2000] como um aspecto mais crucial do modelo estocástico, pois o algoritmo precisa saber o que é uma solução ótima. Para isso, o sistema precisa encontrar no poema os aspectos fonéticos regulares, como exemplo de rima, métrica e aliteração. Estas informações são fornecidas por meio de um dicionário de pronúncia. Uma alternativa possível para identificação da forma fonética ideal é especificando um modelo como entrada, através de uma simbologia, que serve como guia para marcação da solução candidata. Uma outra alternativa é especificar um conjunto das formas, com métrica, rima e aliteração para um poema, alimentando o sistema com conhecimento de tipos de poemas existentes.

O sistema de Manurung, Ritchie e Thompson [2000] também faz verificação linguística e semântica. Sobre a linguística, é feita a verificação lexical, para recompensar o poema pelo uso de palavras consideradas interessantes, marcadas como palavras “poéticas”, correspondentes a avaliação sintática. Referente a semântica, o autor considera abstrato o mecanismo desta avaliação, mas segue o mesmo estilo de avaliação da fonética, especificando a semântica alvo, e marcando a semântica candidata.

Já no trabalho de Rashel e Manurung [2014] os instrumentos de avaliação utilizados foram o questionário online, baseado no teste de Turing. Inicialmente, verificou-se o comportamento do sistema, para determinar como os mesmos são capazes de imitar o comportamento humano na escrita de poemas. Para este teste, foram utilizadas as palavras poéticas das variadas fontes de entrada, evitando a discussão de apenas um tema específico. Alguns destes poemas foram apresentados em um questionário, e avaliados por um público mesclado de indonésios apreciadores de poemas, comunidade científica e pessoas comuns, convidados a identificar se o poema foi escrito por humano ou pelo sistema.

Como resultado dos questionários aplicados por Rashel e Manurung [2014], 74% dos avaliadores identificaram corretamente os poemas escritos por humanos, e 26% foram identificados equivocadamente como poemas escritos pelo computador. Já os poemas gerados conforme o conjunto completo de restrições, 57% dos avaliadores erraram as suas identificações, considerando os poemas como criações humanas, mostrando um bom resultado de pesquisa. Um outro bloco de avaliações correspondeu a poemas gerados e totalmente guiados pelo sistema, dos quais 35% foram identificados erroneamente como poemas de autoria humana.

A avaliação do sistema BertsoBOT também foi feita através do teste de Turing, quando Agirrezabal *et. al.* [2013] contaram com dois linguistas que não tinham participado do projeto, para identificarem as estrofes que foram criadas por humanos, e as que foram criadas pelo computador. Tais estrofes foram embaralhadas, totalizando 135, e após três etapas de avaliação, Agirrezabal *et. al.* [2013] chegaram à conclusão de que houve um avanço considerável sobre os resultados da segunda etapa, quando os avaliadores imaginaram que 26% dos versos gerados pela máquina, tinham sido feitos por humanos.

Apesar dos sistemas de Manurung [2004] e de Misztal e Indurkha [2014] serem capazes de validar os resultados obtidos através de um conjunto de dimensões predefinidas, eles consideram propriedades poéticas, gramática e significado. Para avaliarem os resultados obtidos recorreram a avaliadores humanos, da mesma forma que Das e Gamback [2014], Zhang e Lapata [2014] e [2016] aplicaram questionários, porém, cada pesquisador elaborou as suas perguntas de acordo com as saídas do sistema, acrescentando perguntas que refletissem a emoção, o gosto e a compreensão do poema. A escolha pelo melhor soneto gerado pelo sistema de Ghazvininejad *et. al.* [2016] também foi feita por 23 juízes humanos, com base em dez diferentes tópicos para soneto e estrofes.

Para avaliar os poemas gerados pelo sistema Das e Gamback [2014] contaram com três avaliadores, um é estudante da literatura Bengali, outro é jornalista e o terceiro estudante de tecnologia, para testar o sistema com 100 entradas de frases escolhidas por cada um deles. De modo geral, as avaliações foram negativas, sendo que apenas o estudante de tecnologia que atribuiu uma boa pontuação à poética, mas atribuiu pontos inferiores à gramática e ao significado.

Na avaliação do sistema Tra-la-lyrics 2.0, [Oliveira, 2015] também utilizou questionário, por meio de perguntas relacionadas a ritmo, rima, som, gramática, significado, assunto, qualidade geral, pontuados de 1 a 5, considerando 287 pesquisas feitas com 94 letras, tendo uma pontuação média geral de 3,01. O destaque é para canção Alecrim, que obteve nota 4 em todos os aspectos. No futuro Oliveira [2015] pretende que o seu sistema seja capaz de cantar a letra produzida, com o incremento de uma ferramenta de síntese de voz. No seu sistema seguinte Oliveira [2017] acrescentou restrições nos procedimentos aplicados para avaliada idioma, pontuando as dimensões de características poéticas de métrica e rima, a variação da estrutura dos versos e a conexão feita entre a semântica e as palavras fornecidas.

Devido à dificuldade de avaliar sistemas criativos automaticamente, muitos pesquisadores recorreram a um público que pudesse realizar a exaustiva tarefa com base em um conjunto de dimensões predefinidas. Porém, existem algumas implementações algorítmicas capazes de avaliar os sistemas de geração de poemas ou os seus resultados. Os especialistas automáticos foram explorados por Gervás [2013] e Misztal e Indurkha [2014] como parte do processo para avaliar métrica e rima, sendo um grupo de indivíduos é formado com base em informações iniciais de semântica e fonética alvo. Cada poema recebe uma pontuação, e na etapa de evolução é selecionado o que tiver maior pontuação para reprodução, alcançando como resultado as suas melhores versões.

Como forma de comparação entre os elementos utilizados por alguns pesquisadores na geração de poemas computacionais, a Tabela 2.18 apresenta uma implementação de cada idioma, com respectivos tipos de poemas, fontes de dados para entrada no sistema, as técnicas do domínio da inteligência artificial, e as formas de avaliação para cada sistema. Os pesquisadores fizeram opção pelo idioma local, sendo que Oliveira *et. al.* [2017] fizeram adaptações no sistema PoeTryMe para mais dois idiomas, sendo algumas linguagens mais ou menos fonéticas.

Quanto aos dados de entrada, a escolha por poemas permite que os versos sejam recombinaados para gerar novos poemas. Quanto a escolha por textos, é feita uma busca por frases que atendam a parametrização estabelecida pelo usuário. Uma outra forma de entrada de dados é por meio de palavras aleatórias que podem ser recombinaadas para gerar o poema de acordo com as regras gramaticais. De cada tipo de dado de entrada no sistema é gerado um tipo de poema através de variadas técnicas do domínio da inteligência artificial. Todos os poemas foram avaliados com

pontuação para cada requisito, uns são avaliados pelo próprio sistema, outros por pessoas e alguns foram avaliados pelos dois meios.

Tabela 2.18: Comparação entre trabalhos dos autores.

Pesquisador	Idiomas	Dados Entrada	Dados Saída	Técnica	Avaliação
Agirrezabal, 2013	basco	Poemas	pequeno de oito	Geração de Linguagem Natural	teste de Turing
Chrismartin, Tobing, e Manurung, 2015	inglês	artigos de notícias	quadras	Algoritmo evolutivo	computador
Das e Gambäck, 2014	bengali	Poemas	poemas infantis	Geração de Linguagem Natural	computador e humano
Ghazvininejad <i>et. al.</i> , 2016	espanhol	sinônimos de dicionário	quinteto e soneto	Rede Neural	humano
Gervás, 2000	espanhol	sonetos clássicos	romance e soneto	RBC	computador e humano
Gervás, 2001	espanhol	mensagem em prosa e poemas	poesia formal	RBC	computador e humano
Gervás, 2016	espanhol	poemas combinados com artigos mexicanos	poesia formal	RBC	computador
Manurung, Ritchie e Thompson, 2000	inglês	palavras aleatórias	poema regular	Modelo Estocástico Hill Climbing	computador
Manurung, 2004	indonésio	informações do usuário	poema regular	Algoritmo evolutivo	computador
Manurung, 2014	indonésio	artigo de notícias	poema clássico	Algoritmo evolutivo	teste de Turing
Manurung, 2015	indonésio	artigo de notícias	poema clássico	Tabela de geração	teste de Turing
Misztal e Indurkha, 2014	indonésio	blogs	poema clássico	quadro negro	computador
Oliveira, 2012	português europeu	rede social Twitter	haiku, soneto	Tabela de geração	computador e humano
Oliveira <i>et. al.</i> , 2017	português europeu, espanhol e inglês	rede social Twitter	haiku, soneto	Geração de Linguagem Natural	computador
Rashel e Manurung, 2014	espanhol	letras de música lírica	poema regular	Geração de Linguagem Natural	teste de Turing
Yan, 2016	inglês	tópicos, temas.	quadras	Deep Neural Network	Humano
Zhang e Lapata, 2016	chinês	poemas online	poemas clássicos	Deep Neural Network e Recurrent Neural Network	Humano

Capítulo 3

O Sistema

“Conte-me e eu esqueço. Mostre-me e eu apenas me lembro. Envolve-me e eu compreendo.”

-- Confúcio

Este capítulo apresenta o sistema PROPOE (*Prose to Poem Automatic Generator*), considerado como uma ferramenta capaz de receber sentenças metrificadas e gerar poemas, com variados tipos de rima e ritmos. Desta forma, são descritas a abordagem e técnica utilizada na implementação do mesmo, para alcance das melhores combinações entre versos, e as rotinas utilizadas na sua parametrização. Além disso, inicialmente será citado o processamento do MIVES, pois este sistema é responsável por exportar as estruturas de versificação dos textos de obras literárias, a serem explorados pelo sistema PROPOE.

No processo de geração de poemas, foram consideradas as regras estabelecidas para o português do Brasil, encontradas nas obras de [Goldstein, 2006] e [Silva, 2014] que serviram como base de pesquisa. No processo de implementação do sistema, são considerados algumas das abordagens aplicados em outras implementações, como exemplo da avaliação final por pontuação que o poema pode receber, de acordo com a identificação de algum padrão considerado ideal em poemas regulares, sob a visão das formas de rima identificadas nas palavras que compõem os versos, bem como o ritmo.

Os primeiros testes do PROPOE foram realizados pelo próprio sistema, na obra de “Os Sertões”, devido a versificação das sentenças já ter sido realizada por pesquisadores desta obra, fornecendo informações relevantes sobre as suas estruturas métricas, para posterior combinação das mesmas na identificação das rimas e ritmos. Porém, esta não é a única obra que o PROPOE pode trabalhar, pois o sistema é capaz de gerar poemas a partir de qualquer obra que tenha sido processada pelo MIVES, que fornece sentenças com estruturas métricas.

Uma visão geral do PROPOE é apresentada na Figura 3.1 com as etapas do processo. O sistema PROPOE carrega um *corpus* de segmentos frásicos que tenham sido processados no MIVES, com as suas estruturas versificadas. Tais estruturas são classificadas pelo número de sílabas poéticas, posicionamento das sílabas tônicas, e sentenças escandidas. A partir destes dados, o PROPOE identifica a última sílaba da palavra de cada sentença, fazendo o agrupamento de cada sílaba com as sentenças terminadas com a mesma sílaba. Essa lista de sentenças por sílaba passa por um filtro para que seja reduzido o número de sentenças candidatas a serem combinadas no momento da montagem do poema. Tanto a filtragem quanto a montagem são guiadas por parâmetros.

As sentenças são filtradas pela escolha da métrica, do número de estrofe, do número de verso e pelas sílabas a serem combinadas. Com esse filtro é possível gerar uma sublista de sentenças, reduzindo o número de possibilidades de combinações entre as sentenças candidatas a compor do poema.

A estruturação do poema ocorre pelo critério de montagem, estabelecido através da escolha do tipo de rima, da similaridade, e da comparação de versos. O tipo de rima define dentre as sentenças que combinam, em que ordem elas irão aparecer no poema. A similaridade tanto determina a exibição dos versos pela igualdade do esquema rítmico, quanto determina a exibição pela igualdade entre sentenças de mesmo posicionamento de sílaba tônica.

Os poemas gerados são avaliados quanto a semelhança entre as letras das palavras que rimam, a coincidência entre as sílabas tônicas dos versos, a semelhança entre os posicionamentos das sílabas tônicas das palavras que rimam, além do esquema rítmico entre os versos. Esta avaliação é feita por meio de uma pontuação estabelecida para de cada critério e do total representando a soma das pontuações.

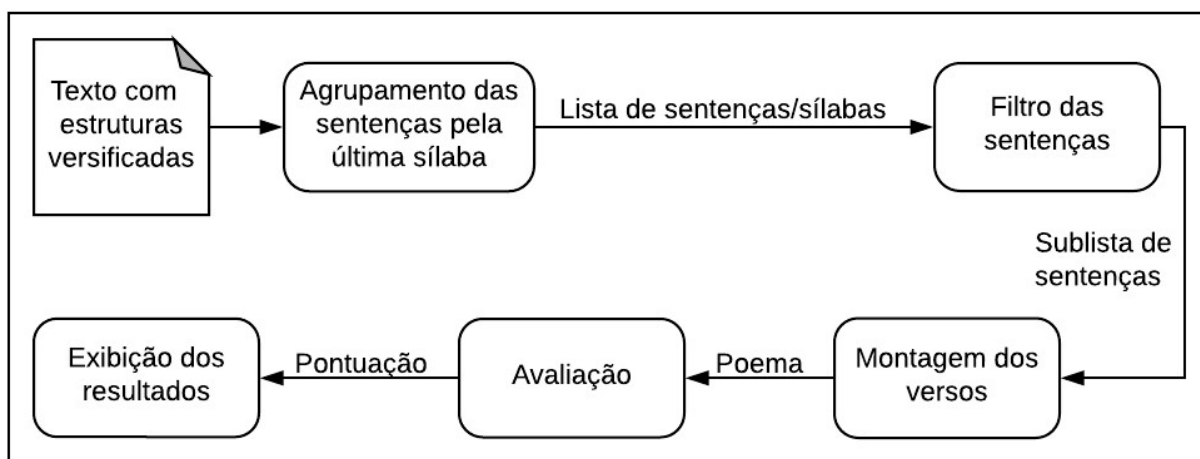


Figura 3.1: Etapas de processo do sistema.

3.1 Texto com estruturas versificadas

O texto original de obras literárias em prosa, passa inicialmente pelo processamento do sistema MIVES [Carvalho; Loula E Queiroz, 2016] que realiza a escansão das sentenças, e identifica as suas respectivas métricas, por meio da extração de todos os segmentos do texto, separados por um sinal de parágrafo. Na próxima etapa do MIVES, ocorre a separação silábica ortográfica de cada palavra das sentenças, bem como a escansão e contagem das sílabas poéticas, o posicionamento das sílabas tônicas e a classificação da métrica dos versos, possibilitando a exportação de um arquivo contendo todas as sentenças identificadas sequencialmente na obra literária, e as suas características.

O formato das sentenças e suas características exportadas pelo MIVES podem ser vistas na Tabela 3.1, que expõe apenas como exemplo dois dos seus segmentos, de toda a obra literária de “Os Sertões”, com respectivos números de sílabas, posição das sílabas tônicas, e as sentenças escandidas, ou seja, separadas em sílabas poéticas, com marcação das sílabas tônicas. Estes dados podem ser exportados em formato XML, sendo essenciais para o processamento do

PROPOE que ao receber este arquivo como entrada no sistema exibe também a frequência de sentenças por sílaba e a frequência de sentenças por ritmo.

Tabela 3.1: Estruturas versificadas extraídos do MIVES.

```

<exportacao.Sentenca>
<segmento>Hipóteses sobre a sua gênese.</segmento>
<estruturaDeVersificacao>
  <exportacao.EstruturaVersificacao>
    <numeroDeSilabas>10</numeroDeSilabas>
    <posicaoDasTonicas>2 5 8 10</posicaoDasTonicas>
    <sentecaEscandida>Hi/p#ó/te/ses/ s#o/bre/ a/ s#u/a/ g#ê/ne/se.</sentecaEscandida>
  </exportacao.EstruturaVersificacao>
</estruturaDeVersificacao>
</exportacao.Sentenca>
<exportacao.Sentenca>
<segmento>A região continua alpestre.</segmento>
<estruturaDeVersificacao>
  <exportacao.EstruturaVersificacao>
    <numeroDeSilabas>10</numeroDeSilabas>
    <posicaoDasTonicas>4 7 10</posicaoDasTonicas>
    <sentecaEscandida>A/ re/gi/#ão/ con/ti/n#u/a/ al/p#es/tre.</sentecaEscandida>
  </exportacao.EstruturaVersificacao>
</estruturaDeVersificacao>
</exportacao.Sentenca>

```

3.2 Organização de Sentenças por Sílaba

Uma das formas de identificar rimas em um poema é através da última sílaba da última palavra de cada verso, sendo este classificado como sentença, antes de fazer parte de um poema. Como exemplo da entrada inicial, é apresentada na Tabela 3.2 uma mera ilustração de algumas das sentenças identificadas no arquivo XML da obra de “Os Sertões”.

Tabela 3.2: Sentenças escandidas do MIVES.

Escansão: Es/t#a/va/ blo/qu/e/#a/do/ Ca/n#u/dos. Metrica: Decassílabo Número de Silabas: 10 Posição das Tônicas: 2 7 10
Escansão: Per/d#i/am-/se/ nos/ ca/s#e/bres/ u/n#i/dos. Metrica: Decassílabo Número de Silabas: 10 Posição das Tônicas: 2 7 10

As sentenças anteriormente exibidas são facilmente identificadas pela marcação de um ponto ao final de cada uma delas. Além disso, elas são separadas em sílabas poéticas marcadas por barras, e as sílabas fortes, são marcadas por um símbolo de cerquilha para informar o

posicionamento das sílabas tônicas. Nesta lista, são identificadas três sentenças da sílaba “dos”, três sentenças da sílaba “mo”, e duas sentenças da sílaba “va”.

Para estruturar as sentenças no PROPOE, todas as últimas palavras de cada sentença foram quebradas em sílabas, com o intuito de agrupar as sentenças e demais dados relacionados, por sílaba, fazendo a seleção de um número de sentenças de determinada sílaba. A Tabela 3.3 mostra como tais dados são organizados sequencialmente, na ordem em que se encontram as sentenças no arquivo fonte, agrupando-as às suas correspondentes sílabas finais.

Tabela 3.3: Algumas sentenças do arquivo XML.

Chave	Valor
Dos	Es/t#a/va/ blo/qu/e/#a/do/ Ca/n#u/dos. Metrica: Decassílabo Número de Silabas: 10 Posição das Tônicas: 2 7 10 Per/d#i/am—/se/ nos/ ca/s#e/bres/ u/n#i/dos. Metrica: Decassílabo Número de Silabas: 10 Posição das Tônicas: 2 7 10 Ex/pun/g#i/ra O a/rrai/#al/ de in/cr#é/du/los/ E/ t#i/mi/dos. Metrica: Decassílabo Número de Silabas: 10 Posição das Tônicas: 2 7 10
mo	Bus/qu#e/mos/ um/ e/x#em/plo/ #ú/ni/co, O #úl/ti/mo. Metrica: Decassílabo Número de Silabas: 10 Posição das Tônicas: 2 6 8 10 Re/fr#ão—/ con/si/d#e/ram—/no/ ga/li/c#is/mo. Metrica: Hendecassílabo Número de Silabas: 12 Posição das Tônicas: 2 6 11 Am/pl#i/am/ a/ in/flu/#ên/ci/a/ do #úl/ti/mo. Metrica: Decassílabo Número de Silabas: 10 Posição das Tônicas: 2 7 10
va	For/m#a/vam-/se/ Os/ c#or/pos/ de/ re/s#er/va. Metrica: Eneassílabo Número de Silabas: 9 Posição das Tônicas: 2 5 9 A/ an/t#i/ga/ ca/p#e/la/ n#ão/ bas/t#a/va. Metrica: Decassílabo Número de Silabas: 10 Posição das Tônicas: 3 6 8 10

A partir da lista contendo as estruturas de versificação, uma nova lista é formada contendo as últimas palavras de cada sentença. Tais palavras sofrem uma quebra em sílabas, e são

organizadas em uma estrutura de hash, fazendo uma relação entre sílabas que correspondem à chave, e as respectivas sentenças em associação com suas escansões, métricas e esquemas rítmicos, que são os valores de cada chave.

3.3 Filtro das Sentenças

Como o arquivo recebido pelo sistema contém sentenças métricas de toda obra literária escolhida, e com variadas particularidades nas suas métricas, houve a necessidade de estabelecer o critério de filtragem para reduzir o número de sentenças a serem manipuladas a cada interação. Este filtro permite o carregamento em uma lista, apenas das sentenças correspondentes à parametrização escolhida, para realizar a combinação entre um número reduzido delas, minimizando o custo computacional. Pois considerando que o número de sentenças métricas da obra “Os Sertões” seja de 1.428 sentenças, para gerar uma simples quadra de rima mista, é possível realizar 4.158.271.385.856 combinações.

O filtro inicia pela métrica, quando o sistema considera a escolha de quaisquer das métricas existentes, de acordo com o número de sílabas poéticas desejado para cada verso. As métricas disponíveis no sistema, são de uma a treze sílabas, porém, o MIVES faz identificação de seis a treze sílabas nas sentenças, e dessa forma, na obra de “Os Sertões” tem sentenças nestas métricas para gerar poemas. O PROPOE também permite a escolha de não filtrar por métrica, e assim utilizar por todas as métricas para formar um poema heterométrico. Seja qual for a escolha métrica, o sistema busca por uma lista de sentenças com respectivas estruturas versificadas que podem atender à solicitação.

Na sequência, devem ser escolhidos o número de estrofes e o número de versos que se deseja para o poema. Este parâmetro controla o processo de montagem mas também é utilizado na filtragem. A consequência desta escolha é a filtragem de sentenças por sílaba, cujas sentenças relacionadas são capazes de atender à solicitação. Isso evita que haja repetição de uma mesma sentença em um mesmo poema, por falta de quantidade de sentenças diferentes. Assim, caso seja escolhido um número de estrofes e versos, cuja sílaba não possua sentença suficiente para atender à solicitação, esta não é exibida.

Para filtragem de sentenças por sílaba, duas ou quatro sílabas podem ser escolhidas tanto manualmente através da relação exibida, quanto pelo sistema de forma aleatória. Na escolha aleatória, o próprio sistema se encarrega de realizar o sorteio dentre as sílabas disponíveis, sempre com base nas sentenças existentes para atender à solicitação. O poema será gerado na etapa seguinte com base na rima de duas ou quatro sílabas, considerando duas das sílabas para cada estrofe. A filtragem é feita em cascata, e vai acontecendo até que o poema seja gerado, reduzindo o número de sentenças para alcance do produto final, a ser exibido com base na etapa de montagem do poema.

3.4 Montagem do poema

A montagem do poema corresponde a estruturação de estrofes em versos, a partir de uma variada lista de sentenças que tenham sido filtradas. O problema a ser tratado corresponde a escolha entre as sentenças candidatas e a determinação da ordem em que as mesmas devem aparecer. Uma estratégia possível seria testar diversas combinações para obter o melhor ritmo

no poema como um todo, ou seja, aplicar uma estratégia de otimização global no poema. Porém, dado o grande número de sentenças possíveis, mesmo após a filtragem, avaliar todas as combinações teria ainda um alto custo computacional.

Estratégia gulosa

Assim, optou-se por uma estratégia gulosa, seguindo um paradigma divisão e conquista, ou seja, é feita a fragmentação do problema a ser resolvido em partes menores, e em seguida as soluções obtidas são combinadas para alcance da solução geral. O algoritmo guloso, também conhecido como ganancioso, é uma técnica com o propósito de solucionar problemas de otimização combinatória, fazendo a escolha que aparentemente é a mais adequada para o momento, gerando a expectativa de que a solução final será a melhor possível. A decisão tomada pelo algoritmo é sempre definitiva, e com base nas informações disponíveis na iteração corrente. Por isso, as consequências futuras são ignoradas. A Figura 3.2 apresenta o fluxograma para montagem do poema.

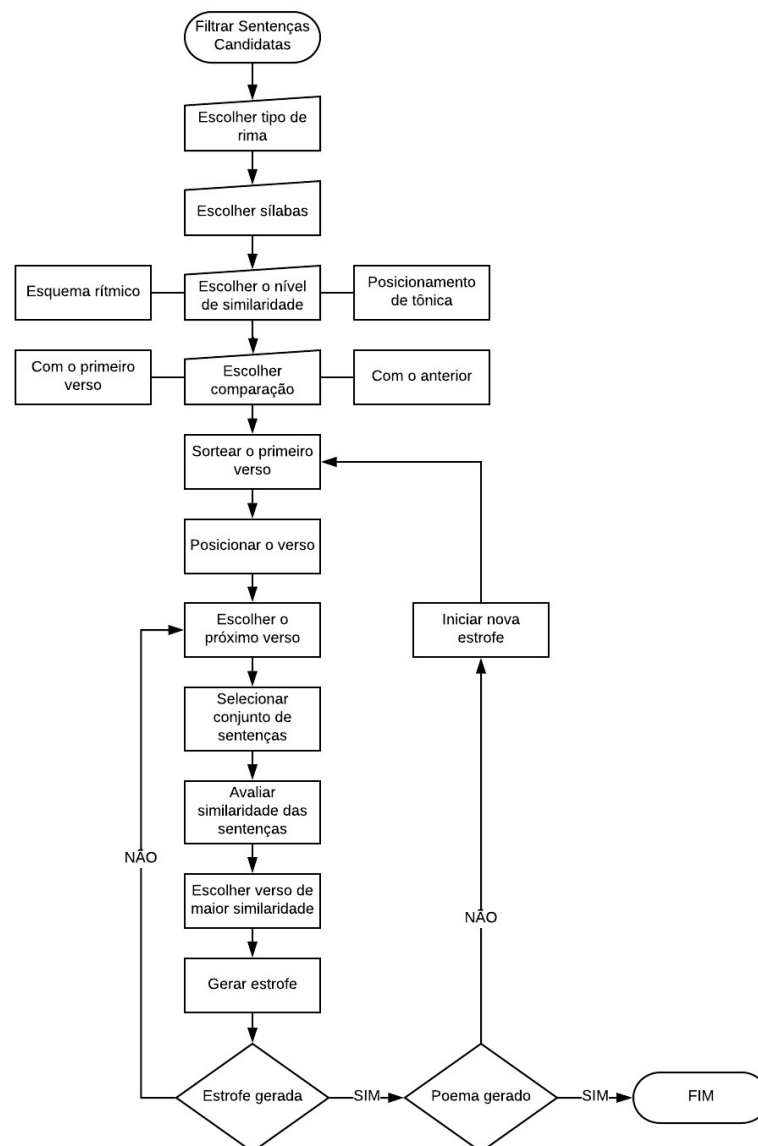


Figura 3.2: Montagem do Poema.

Dessa forma, é possível reduzir o problema de montagem de um poema, em um problema de montagem de estrofes, em que a decisão sobre a escolha de cada verso é feita usando um único verso já posicionado como referência. Esta é uma estratégia então de otimização local, comparando pares de versos, ou seja, versos candidatos são escolhidos comparando com um verso de referência já escolhido. Com isso, um poema será uma concatenação de estrofes montadas por esta estratégia.

Formação de Estrofe

A formação de cada estrofe é feita com base na escolha do tipo de rima, número de estrofes, número de versos, e pelo critério de comparação dos versos, além da escolha de sílabas e lista de sentenças obtidas da filtragem. Para isso, cada estrofe é gerada de forma independente, sendo o primeiro verso escolhido aleatoriamente dentre os disponíveis após filtragem, e os demais versos de cada estrofe escolhidos segundo o critério de comparação. O número de estrofes escolhido em conjunto com a escolha de duas ou quatro sílabas, determinam o posicionamento das sílabas aos versos, alocando sentenças relacionadas conforme sequência de sílabas escolhidas.

Escolha das sílabas

Ao escolher duas sílabas, conforme Tabela 3.4, as estrofes são sempre montadas, considerando a primeira e segunda sílabas, mas com escolha de quatro sílabas, as estrofes ímpares são montadas com sequência de primeira e segunda sílabas, e as estrofes pares, com sequência de terceira e quarta sílabas.

Tabela 3.4: Quantidade de sílabas no poema.

Poema 1 - Duas sílabas	Poema 2 - Quatro sílabas
Tudo aquilo era uma ficção geográfica. Esta escala foi providencial. Di-lo uma comparação histórica. Mal atentaram para o arraial. E falseavam a peleja franca. Contorneavam o arraial. De repente, uma variante trágica. Caía feito um rastilho no arraial.	Acabaram-se as munições de boca. Caía feito um rastilho no arraial. Quedaram na defensiva franca. Esta escala foi providencial. Os assaltantes eram repelidos. Ambas lhe são abstrações inacessíveis. À noite tiroteios cerrados. Os demais prosseguiram impassíveis. Tudo aquilo era uma ficção geográfica. Mal atentaram para o arraial. Di-lo uma comparação histórica. No outro dia atingiram o arraial.

Escolha da rima

A escolha da rima determina o posicionamento dos versos em cada estrofe do poema. Nos poemas apresentados da Tabela 3.4, houve a escolha pelo tipo de rima cruzada, com duas representações de estrofes. Para o poema 1 foi feita a escolha das sentenças das sílabas finais “ca” e “al”. Para o poema 2, as estrofes ímpares são formadas por sentenças com sílabas finais “ca”, “al”, e a estrofe par é gerada com as sentenças de sílabas finais “dos” e “veis”. A sequência de sílabas escolhida é associada a um tipo de letra apenas para demonstração de esquema rítmico, considerando a sequência de A, B, C e D, e isso classifica o tipo de rima quanto a posição do verso na estrofe.

Além da rima cruzada, o PROPOE também gera poemas em quadras de rima emparelhada, mantendo a rima numa sequência de cada dois versos, como AABB; de rima interpolada, rimando o primeiro verso com o quarto, e o segundo com o terceiro, formando ABBA; e de rima mista, que exhibe um poema formado sem imposição de ordenação pré-definida dos versos na estrofe. Na rima mista, cada verso é sorteado tanto para A quanto para B, e por isso aumenta o número de combinações entre as rimas, podendo formar AAAA, AAAB, AABA, ABAA, BAAA, BBBB, BBBA, BBAB, ou qualquer combinação das rimas anteriormente descritas, que pode ser vista na Tabela 3.5.

Tabela 3.5: Rima mista.

Esta impõe-se, tenaz e inflexível.
 Esta tendência é expressiva.
 Mas de repente o tumulto cessava.
 Não havia errar o alvo imóvel.
 E a tarefa não é insuperável.
 O antagonismo era inevitável.
 Previram a derrota inevitável.
 Ele ali subia e pregava.

O poema de rima mista apresentado na Tabela 3.5 possui sentenças decassílabas, escolhidas de forma aleatória. A primeira estrofe possui rima na forma ABBA, sugerindo rima interpolada, mas não considerada como este tipo de rima, devido a sequência apresentada na segunda estrofe de AAAB, oferecendo uma nova forma para o poema, que é considerado como rima mista. Esta situação pode ocorrer, devido a aleatoriedade na escolha de cada verso, podendo gerar poemas com variadas formas tanto na primeira estrofe, quanto nas estrofes subsequentes.

Considerando poemas de rimas cruzada, emparelhada ou interpolada, um fator que pode descaracterizá-las é o número de versos por estrofe, pois apesar do PROPOE gerar poemas com os mais variados números de versos, a rima é baseada em quartetos, e caso esta quantidade seja menor, o padrão será desconsiderado e o poema não segue o padrão de rima definido para cada estrofe. Assim, caso a quantidade de versos for maior que 4, o padrão será repetido total ou parcialmente, de forma similar.

A Tabela 3.6 apresenta as rimas cruzada, emparelhada e interpolada com duas ou quatro sílabas, demonstrando as combinações bem sequenciadas, conforme o padrão estabelecido para cada tipo de rima. Porém, tal padrão é quebrado nas estruturas de três versos, afinal, a cada nova estrofe inicia-se um novo padrão para sequência e não uma continuação da estrofe anterior.

Tabela 3.6: Posicionamento dos versos.

Rima	Núm. Estrofe	Num. Sílabas	2 versos	3 versos	4 versos
Cruzada	1	2	AB	ABA	ABAB
Emparelhada	1	2	AB	AAB	AABB
Interpolada	1	2	AB	ABB	ABBA
Cruzada	2	4	AB CD	ABA CDC	ABAB CD CD
Emparelhada	2	4	AB CD	AAB CCD	AABB CCDD
Interpolada	2	4	AB CD	ABB CDD	ABBA CDDC

Além das combinações descritas na Tabela 3.6, para cada sílaba A, B, C ou D, ainda existem diversas sentenças possíveis de serem utilizadas para cada sílaba. A primeira sentença de cada estrofe é escolhida de forma aleatória como semente inicial da estrofe, e as seguintes, são escolhidas por similaridade tendo um verso como referência. O verso de referência pode ser o primeiro da estrofe, ou o verso imediatamente anterior ao verso a ser escolhido, conforme parâmetro da montagem, escolhido pelo usuário. Tais critérios de comparação independem do tipo de rima escolhido.

Nível de Similaridade

Esta etapa evidencia a coincidência dos sons entre os versos. Na busca por similaridade acontece a escolha pelo verso que tenha um esquema rítmico similar ou acentuação da última tônica mais similar ao verso de referência, e a escolha entre estas duas possibilidades é um parâmetro de escolha do usuário. O esquema rítmico corresponde ao posicionamento das sílabas fortes na sentença, e a acentuação corresponde à coincidência da posição da última sílaba tônica nas palavras que rimam. Mas pode haver empates nessa busca, caso exista mais de um verso com o mesmo esquema rítmico (ou acentuação). Para fazer a escolha entre um deles, o sistema decide escolher aquele que tenha também a melhor acentuação (ou esquema rítmico), que serve como critério de desempate.

Para atender a uma das possibilidades dentre esquema rítmico ou acentuação, o sistema ainda deve considerar se os versos devem ser comparados com o primeiro verso de cada estrofe, ou tal comparação deve ser feita entre o verso anterior e o seu sucessor. Dessa forma, a Tabela 3.7 apresenta dois poemas gerados em forma de quadra, exibidos como sentenças escandidas para melhor percepção do esquema rítmico, sendo que em um poema, o verso referência é o primeiro, e o outro, tem um primeiro verso como semente, mas compara o esquema rítmico de cada verso com anterior.

Tabela 3.7: Poemas com diferentes combinações de versos.

Similaridade	Sentenças Escandidas	Esquema Rítmico	Acentuação
Comparados com o primeiro	N#ão/ #e/ram/ si/ti/#an/tes,/ #e/ram/ pr#e/sos.	1 2 6 8 10	Grave
	Vo/ze/#a/vam/ a/cla/ma/ç#ões/ E a/pl#au/sos.	3 8 10	grave
	Ca/#ia/ f#ei/to um/ ras/t#i/lho/ no a/rrai/#al.	2 3 6 10	aguda
	No #ou/tro/ d#ia/ a/tin/g#i/ram/ O a/rrai/#al.	1 3 6	aguda
Comparados com o anterior	A/ss#al/tam-/no;/ a/f#e/rram-/no;/ ju/g#u/lam-/no.	2 6 10	Esdrúxula
	Os/ ja/g#un/ços/ à/ p#or/ta/ a/ssal/t#a/vam-/no.	3 6 10	esdrúxula
	A/ ex/tr#e/ma/ d#or/ #e/ra/ a extrema-/un/ç#ão.	3 5 6 10	aguda
	V#i/mos/ qu#an/tos/ en/tr#a/ram/ Em/ a/ç#ão.	1 3 6 10	aguda

Os poemas da Tabela 3.7 São formados por dois quartetos cada, mas com critérios diferentes para comparação entre as similaridades. No primeiro poema, o verso que serve como referência tem o esquema rítmico 1 2 6 8 10, e os demais versos são carregados de acordo com estes ritmos. No segundo poema, o primeiro verso da estrofe é referência apenas para o próximo verso, e assim sucessivamente, ou seja, o verso anterior é sempre referência do seu sucessor. A busca pelo verso que tenha similaridade com o verso de referência, pode ocasionar em empate, por ter mais de um com o mesmo esquema rítmico. Esta situação requer a escolha por um verso, quando é eleito aquele que tenha também a mesma semelhança de acentuação.

Na busca pela similaridade de esquema rítmico, o PROPOE realiza o cálculo de similaridade com base no coeficiente de Jaccard Equação (1), que permite qualificar a similaridade entre conjuntos, das variáveis A e B, correspondendo respectivamente aos esquemas rítmicos de duas comparações. Tal cálculo é feito dividindo a intersecção entre os posicionamentos das sílabas tônicas dos versos comparados pela união dos posicionamentos das sílabas tônicas.

A execução deste cálculo permite a análise do esquema rítmico que tanto pode ser feito comparando os versos de cada estrofe com o primeiro, como comparando próximo verso com o anterior. Dessa forma, cada comparação é feita pela mesma equação, em que maior valor encontrado para cada combinação realizada resulta em 1, e o verso escolhido para compor o poema é o de maior valor de similaridade.

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{1}$$

Na busca pela similaridade de acentuação da última tônica, o PROPOE obtém a lista de versos disponíveis, e filtra os que possuem a mesma acentuação no final dos versos comparados. Assim, a acentuação do verso corresponde à posição na palavra da última sílaba do verso pronunciada fortemente, podendo ser aguda, grave ou esdrúxula. Na existência de mais de um verso candidato com a mesma acentuação, a escolha é feita com base no segundo critério para escolher um novo verso que tenha o esquema rítmico mais próximo ao verso tido como referência.

O poema pode ser formado por versos isométricos ou heterométricos, e sua filtragem afeta a montagem, pois é provável a existência de um mesmo verso com métricas diferentes. Dessa forma, ao gerar um verso heterométrico, o algoritmo monta uma nova lista de segmentos, com as suas respectivas sentenças escandidas, e marca pelo segmento rítmico quando uma delas é utilizada para gerar o poema, impedindo a existência de repetição.

A Tabela 3.8 apresenta o exemplo de um mesmo verso “O antagonismo era inevitável”, escandido com variação na sua métrica, e conseqüentemente com diferentes marcações de tônicas, - marcadas pelo símbolo de cerquilha -, sugerindo três possibilidades rítmicas. O fenômeno ocorre por meio da elisão, que no caso da métrica decassílabo, une as sílabas poéticas “O” e “an”, por serem vogais átonas, e na situação da métrica do tipo hendecassílabo, une as sílabas “ra” e “i”, pelo mesmo motivo anterior.

Tabela 3.8: Poemas com diferentes combinações de versos.

Segmento	Sentença escandida	Marcação tônica
O antagonismo era inevitável.	O an/ta/go/n#is/mo/ #e/ra i/ne/vi/t#á/vel.	4 6 10
	O/ an/ta/go/n#is/mo/ #e/ra i/ne/vi/t#á/vel.	5 7 11
	O/ an/ta/go/n#is/mo/ #e/ra/ i/ne/vi/t#á/ve.	5 7 12

Por existir mais de um parâmetro na montagem do poema, a técnica utilizada para escolha da melhor sentença no momento da geração do mesmo, é a do algoritmo guloso, que permite a otimização combinatória local com base nos parâmetros e critérios definidos. A aplicação desta técnica, consiste em encontrar um determinado elemento de valor ótimo, em um conjunto finito de possibilidades. Visto que as sentenças de entrada no PROPOE formam este tipo de conjunto,

e os seus elementos podem ser combinados, de acordo com os seus esquemas rítmicos, então, consideramos como uma solução adequada para este cenário.

3.5 Avaliação do poema

Após a montagem do poema, é preciso realizar uma avaliação geral do mesmo produzido pelo PROPOE particularmente, segundo padrões de ritmo para poemas regulares. Para isso é feita uma avaliação com base em critérios, que calculam as pontuações, normalizando valores na escala entre zero e um, em cada avaliação. A maior pontuação que o poema pode receber é 4,0, pois cada elemento avaliado no poema recebe 1,0 como maior pontuação.

Os critérios estabelecidos para avaliação do poema são relacionados à similaridade de posição de todas as tônicas, coincidência das sílabas nas posições de tônicas, soma de sequências iguais de última tônica, e fonética entre as palavras que rimam. Cada um destes critérios tem a sua contribuição na avaliação do poema. Os seus valores são normalizados para o intervalo de zero a um, e a avaliação final do poema é a soma das avaliações por cada critério.

Esquema rítmico

No critério de similaridade de posição de tônicas, será avaliado o esquema rítmico geral do poema que contribui na avaliação do mesmo, pela importância de proporcionar um efeito sonoro de acordo com as suas sílabas átonas e tônicas. Dessa forma, é feita a comparação do posicionamento das sílabas tônicas de cada verso com o verso de referência, de maneira análoga ao cálculo de similaridade de esquema rítmico efetuado na montagem. Em seguida, é feita a soma dos coeficientes encontrados, e para normalizar este valor, é calculada a média da divisão, pelo número de comparações realizadas.

A Tabela 3.9 apresenta um poema de rima emparelhada, com versos contendo ritmos semelhantes, assemelhados ao primeiro verso de cada estrofe, sendo que na primeira estrofe tem dois versos de ritmos iguais. Quanto maior a semelhança entre os versos no poema, maior a pontuação do mesmo, porém na segunda estrofe não ocorre um bom nível de similaridade e isso compromete a pontuação do poema, que neste modelo recebe 0,51 na sua pontuação.

Tabela 3.9: Avaliação do esquema rítmico.

Verso	Métrica	Jaccard	Totalização
A/g#o/ra,/ Nem/ #es/te en/g#a/no/ res/t#a/va.	2 5 7 10		Pontuação estrofe 1 $0,75+0,33+0,4 = 1,48$
Por/qu#e/ a/ a/ç#ã/o/ se/ de/lon/g#a/va.	2 5 10	$3/4 = 0,75$	
Ca/#ia/ f#ei/to um/ ras/t#i/lho/ no a/rrai/#al.	2 3 6 10	$2/6 = 0,33$	Pontuação estrofe 2 $0,6+0,6+0,4 = 1,6$
#Es/ta es/c#a/la/ f#oi/ pro/vi/den/ci/#al.	1 3 5 10	$2/5 = 0,4$	
S#o/bre/ Os/ mu/ra/d#ais/ da i/gr#e/ja/ n#o/va.	1 6 8 10		Pontuação poema $(1,48+1,6)/6 = 0,51$
Co/me/ç#o/u a e/ri/g#ir-/se a i/gr#e/ja/ n#o/va.	3 6 8 10	$3/5 = 0,6$	
No #ou/tro/ d#ia/ a/tin/g#i/ram/ O a/rrai/#al.	1 3 6 10	$3/5 = 0,6$	
M#al/ a/ten/t#a/ram/ para/ O a/rra/i/#al.	1 4 10	$2/5 = 0,4$	

No cálculo do poema acima, a primeira combinação é feita entre as métricas 2 5 7 10 e 2 5 10, cujas tônicas iguais estão posicionadas em 2 5 10, por isso temos 3 combinações que correspondem à intersecção, dividido por 4, que corresponde a união dos dois conjuntos de

tônicas, resultando na pontuação 0,75 para primeira combinação. Este cálculo se repete entre demais comparações, e são somados, conforme coluna de totalização. Ao final deste processo, as somas encontradas nas estrofes são divididas por 6, que é o número de comparações que foi realizada.

Tônicas iguais

O efeito sonoro ganha uma evidência ainda maior, caso as sílabas iguais ocupem a mesma posição tônica nos versos, sendo este um outro critério de avaliação apresentado na Tabela 3.10. Feita a identificação das sílabas tônicas, o poema recebe um ponto a cada sílaba igual identificada. O resultado da soma destas pontuações é dividida pelo número de comparações realizadas. De acordo com o exemplo apresentado, as sílabas iguais são “gre”, “no” e “al”. Dessa forma, o poema recebe 4 pontos pelas sílabas semelhantes, que são divididos por 4 comparações realizadas entre as sentenças que rimam, resultando em 1,0.

Tabela 3.10: Avaliação das Tônicas Iguais.

Versos	Métrica	Tônica Igual	Totalização
A/g#o/ra./ Nem/ #es/te en/g#a/no/ res/t#a/va. Por/qu#e/ a/ a/ç#ã/o/ se/ de/lon/g#a/va. Ca/#ia/ f#ei/to um/ ras/t#i/lho/ no a/rrai/#al. #Es/ta es/c#a/la/ f#oi/ pro/vi/den/ci/#al.	2 5 7 10 2 5 10 2 3 6 10 1 3 5 10	[#al]	1
S#o/bre/ Os/ mu/ra/d#ais/ da i/gr#e/ja/ n#o/va. Co/me/ç#o/u a e/ri/g#ir-/se a i/gr#e/ja/ n#o/va. No #ou/tro/ d#ia/ a/tin/g#i/ram/ O a/rrai/#al. M#a/ a/ten/t#a/ram/ para/ O a/rra/i/#al.	1 6 8 10 3 6 8 10 1 3 6 10 1 4 10	[g#re, n#o] [#al]	2 1 4/4 = 1

Acentuação

Um próximo critério de avaliação é a acentuação das últimas palavras, feito com base nas palavras que rimam, ou seja, cada verso é comparado com o verso anterior da mesma rima. Para isso é feito um cálculo, somando as semelhanças do posicionamento de últimas tônicas de últimas palavras entre versos, e dividindo este total pelo número de comparações realizadas, conforme Tabela 3.11.

Tabela 3.11: Avaliação da similaridade por acentuação.

Poema	Acentuação	Total
Este é um rio sem afluentes. Esta experiência é belíssima. Sucedem-se meses e anos ardentes. Quebra-se o encanto de ilusão belíssima.	a/flu/#en/tes. - Grave be/l#i/ssi/ma. - Esdrúxula ar/d#en/tes. - Grave be/l#i/ssi/ma. - Esdrúxula	1 1
Duplicavam as forças assaltantes. A princípio avançou corretíssima. Desceram ruidosamente as vertentes. O aspecto reduzia-lhe a fama.	a/ssal/t#an/tes. - Grave co/rre/t#i/ssi/ma. - Esdrúxula ver/t#en/tes. - Grave f#a/ma. - Grave	1 3/4 = 0,75

O poema da Tabela 3.11 é formado por rima cruzada, comparando as palavras que rimam “afluentes” e “ardentes”, “belíssima” sendo exibida duas vezes, “assaltantes” e “vertentes”

“corretíssima” e “fama”. Destas comparações são identificadas sílabas tônicas na mesma posição, e por isso recebem a mesma acentuação, sendo duas graves, duas esdrúxulas e outras duas graves. Como cada acentuação igual recebe um ponto, o poema então recebe três pontos dividido por quatro comparações, resultando na pontuação 0,75.

Fonética

Um outro critério de avaliação aplicado corresponde à similaridade entre as letras das palavras que rimam no final de cada verso. Assim, duas das palavras que rimam na estrofe são comparadas e cada coincidência de letra é contada. Esta contagem é então dividida pelo tamanho da menor das palavras comparadas. Tal procedimento é aplicado para cada palavra semelhante, gerando uma nova pontuação. A Tabela 3.12 apresenta o exemplo da avaliação de um poema de rima cruzada, de métrica decassílaba, de sílabas poéticas “do” e “va”.

Tabela 3.12: Avaliação da similaridade fonética do poema.

Poema	Avaliação	Total
Reagiu apesar de ferido. Mas de repente o tumulto cessava. Tiroteios durante o dia todo. A antiga capela não bastava.	cessava - bastava - [a, v, a, s] todo - ferido - [o, d]	4/7 2/4
Foi, sem maior exame, aprovado. Agora, nem este engano restava. Nada referia sobre o passado. Sobre os muradais da igreja nova.	nova - restava - [a, v] passado – aprovado - [o, d, a, p]	2/4 4/7 Soma das razões = 2,14 $2,14/4 = 0,53$

Na avaliação de coincidência das letras é considerado o tipo de rima, por isso as palavras apresentadas acima são comparadas de forma alternada. As letras iguais identificadas nas palavras são apresentadas entre colchetes, e devido a verificação ser feita na ordem invertida das palavras, elas são exibidas obedecendo tal ordem.

Para realizar o cálculo, o PROPOE verifica o número de letras iguais entre as palavras que rimam, e divide pelo número de letras da menor palavra. Em seguida, soma estas razões e divide pelo número de comparações realizadas. No exemplo acima, o cálculo é igual a $(4/7+2/4+2/4+4/7)/4$, sendo esta divisão feita pelo número de combinações realizadas, resultando na pontuação de 0,53.

As pontuações de cada critério de avaliação são somadas e apresentados tanto na interface do sistema, quanto na interface de avaliação, quando é feito um detalhamento sobre cada pontuação.

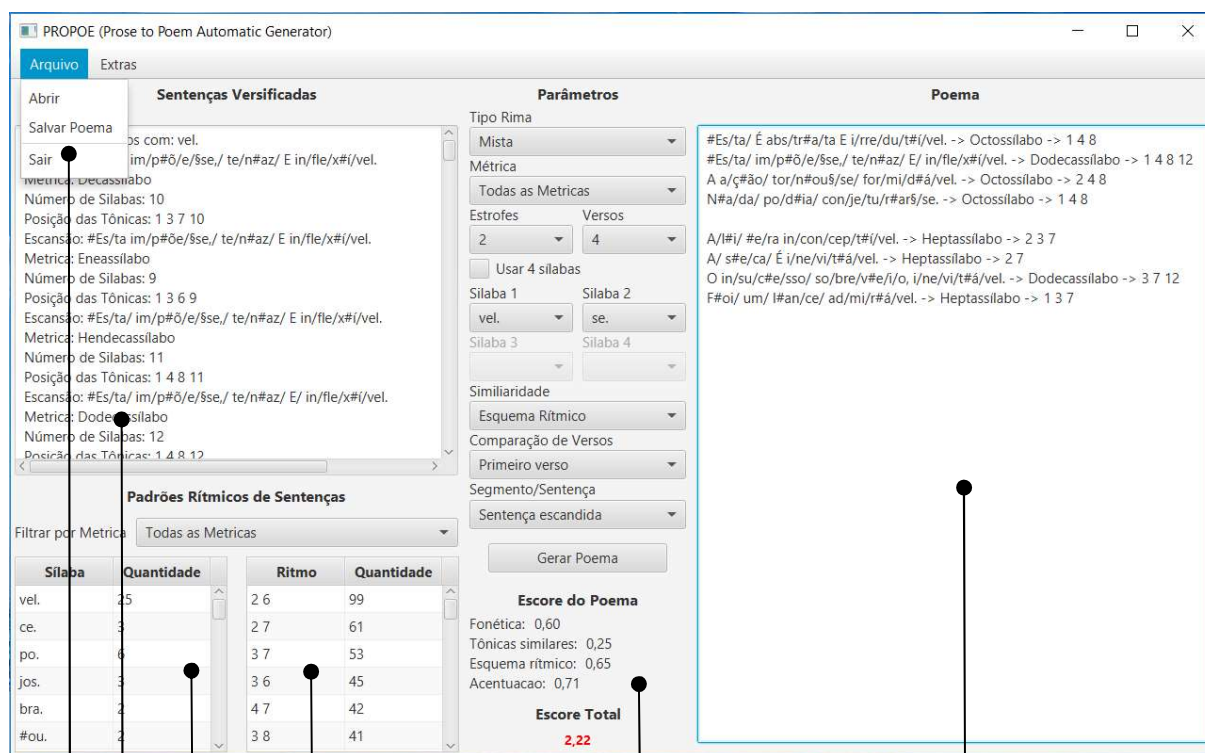
3.6 Interface do Sistema

A interface do sistema é uma consequência da necessidade de analisar a filtragem, montagem e avaliação dos resultados, permitindo ao usuário a visualização das estruturas de versificação que o sistema processa, além de permitir que o usuário realize as suas escolhas de forma

parametrizada, e analise o poema gerado, bem como a avaliação do poema através da pontuação, podendo ser observado na Figura 3.3.

Nesta interface, o PROPOE apresenta em seu menu as opções de abrir o arquivo das sentenças escandidas com as respectivas estruturas versificadas. Uma outra opção é a de salvar o poema em arquivos separados ou acrescentar vários poemas em um só arquivo. A outra opção é a de sair da interface. A frequência de sentenças e frequência de ritmos são carregadas em tabelas, possibilitando a verificação das sílabas que possuem maior número de sentenças, e do esquema rítmico que possui o maior número de sentenças respectivamente. Estas frequências tanto podem ser exibidas por métrica, quanto podem ser apresentadas todas as que servem como entrada no sistema.

Ainda em sua interface, o PROPOE oferece filtros que permitem variadas combinações para gerar o poema. Na área de texto do poema, o mesmo pode ser exibido como tal, ou em forma de sentença escandida, quando os versos aparecem separados em sílabas poéticas, com informações de métrica e marcação no posicionamento das tônicas.



- 1. Menu
- 2. Sentenças Versificadas
- 3. Frequência de Sentenças
- 4. Frequência de ritmos
- 5. Avaliação
- 6. Poema

Figura 3.3: Interface de Visualização de Análise dos Resultados.

Uma vez parametrizado, o sistema possibilita a geração de diversos poemas, avaliados de forma automática, a serem validados pelo usuário. Esta parametrização é referente a critérios de filtro ou de montagem. Como critérios de filtro é possível escolher a métrica de hexassílabo a bárbaro,

e sílabas finais. Apesar da proposta inicial do sistema ser para desenvolver quadras, no critério de montagem são feitas as escolhas pelo número de verso e de estrofes, as variações de tipo de rima cruzada, emparelhada, interpolada ou mista, posicionamento de todas as tônicas, posicionamento da última tônica em cada verso, e definição para comparação de versos.

O poema gerado ao final do processo pode ser exibido em forma de sentença ou em forma de segmento. Em forma de sentença, o poema é exibido com as barras que separam as sílabas poéticas e com os símbolos de cerquilha que marca o posicionamento das sílabas tônicas. Em forma de segmento, o poema é exibido com a ausência dos símbolos informados. A montagem do poema se refere a exibição do mesmo na interface do PROPOE, conforme apresentado na Figura 3.4.

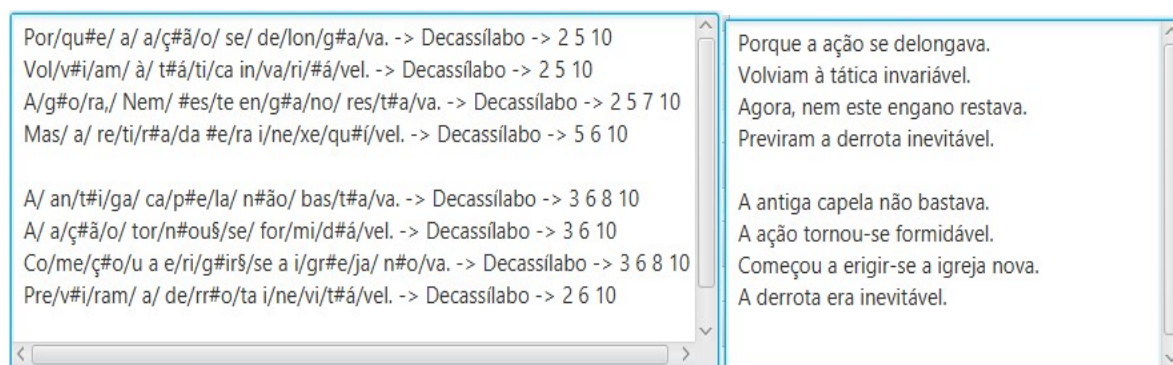


Figura 3.4: Poema exibido em forma de verso e de sentença escandida.

O poema apresentado na Figura 3.4 foi gerado em forma de rima cruzada, com sequência ABAB em cada uma das estrofes, cujas sílabas foram escolhidas de forma aleatória, e o sistema sorteou “va” e “vel”. À esquerda é feita a apresentação do poema em forma de sentença escandida, com informações de métrica, posicionamento de tônica e esquema rítmico. À direita, o poema toma a sua forma comum de exibição. Na forma de sentença escandida, aparece uma letra maiúscula no meio da sentença “A/g#o/ra,/ Nem/ #es/te en/g#a/no/ res/t#a/va.”, bem como símbolo representando um hífen. O sistema realiza o tratamento destas sentenças, convertendo tais caracteres, e no poema tais ocorrências não são exibidas.

As pontuações exibidas na interface principal, também podem ser visualizadas em uma nova tela, que apresenta as avaliações realizadas de forma automática, a ser conferido na Figura 3.4, com detalhamento necessário de amplas informações sobre o poema exibido anteriormente, visto que na tela principal só aparecem os valores das pontuações. Na área relacionada à fonética, são apresentadas as palavras que rimam, e as letras coincidentes entre elas diretamente do vetor, e por isso são exibidas em ordem inversa.

Na área de Posição de Tônica, são exibidas as sentenças escandidas com as marcações, e como o poema foi gerado com a escolha de comparação de todos os versos com o primeiro verso, então o primeiro verso será exibido repetido, e a cada comparação, caso encontre uma sílaba igual, esta é exibida separadamente, e sua pontuação é acrescida também por esta identificação. Na acentuação são exibidas as últimas palavras de cada verso, com a classificação de posição de última tônica do verso.

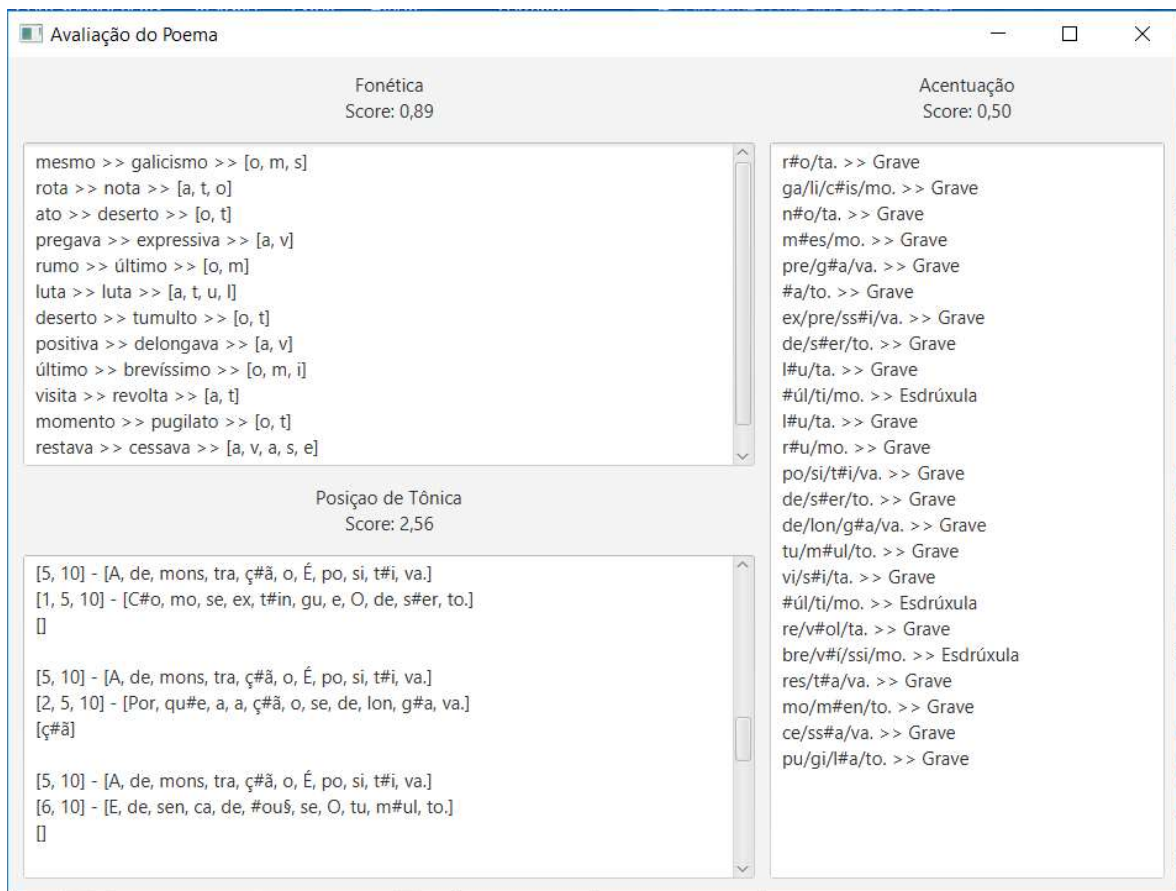


Figura 3.5: Interface de Visualização para Análise dos Resultados.

Capítulo 4

Resultados Alcançados

“A teoria também se converte em graça material uma vez que se apossa dos homens.”

-- Karl Marx

Considerando que o objetivo desta pesquisa é a proposta de um sistema computacional capaz de gerar poemas automáticos a partir de estruturas versificadas de prosa da literatura brasileira, desenvolvemos o PROPOE como uma ferramenta que tem por base a implementação do sistema proposto. Assim, neste capítulo são exibidos os resultados alcançados através do seu funcionamento. Para isso, são apresentadas exemplificações de experimentos realizados, por meio de combinações parametrizadas para gerar poemas a serem avaliados pelo próprio sistema através de pontuações classificatórias de esquema rítmico. Além disso são apresentados os resultados descritivos dos dados nos gráficos de frequência de sentença por sílaba, e de frequência de sentença por esquema rítmico, organizadas pelo PROPOE ao separar as sílabas das sentenças.

Além da obra de “Os Sertões” de Euclides da Cunha, foram testadas também as obras: “Recordações do Escrivão Isaías Caminha” de Lima Barreto, que o autor faz uma crítica à sociedade racista; “Contrastes e Confrontos” de Euclides da Cunha, escrito com base nas publicações feitas pelo autor, reunidas em seu livro; e “Os Lusíadas” de Luís Vaz de Camões. A obra de Camões é poesia épica, composta por 8.816 versos decassílabos, distribuídos em 1.102 estrofes de oito versos cada, sujeitos ao esquema rítmico AB AB AB CC, e por isso considerada como oitava rima camoniana. Apesar da sua estrutura já ser versificada, a obra “Os Lusíadas” também é testada no PROPOE, com o objetivo de realizar novas combinações, com novos esquemas rítmicos. Os experimentos realizados com tais obras são apresentados a seguir.

4.1 Os Sertões

Sobre a obra de “Os Sertões”, o sistema recebe 1.428 sentenças disponíveis, sendo 193 hexassílabos, 257 heptassílabos, 304 Octassílabos, 288 eneassílabos, 291 decassílabos, 300 hendecassílabos, 296 dodecassílabos e 267 bárbaros. Algumas dessas sentenças aparecem na lista com mais de uma métrica, isso ocorre porque existem regras que permitem a junção ou separação de sílabas poéticas.

Resultados descritivos dos dados de entrada

A proposta geral do algoritmo é filtrar os dados e realizar a montagem do poema. Para isso, o sistema seleciona os versos, conforme parametrização, para gerar cada uma das estrofes. O maior número de sentenças identificadas na obra de “Os Sertões” finalizam com palavras cujas últimas sílabas são “te” e “to”, candidatas para gerar o poema. Apesar destas sílabas apresentarem 68 e 57 sentenças respectivamente, na filtragem dos decassílabos são apresentadas 21 sentenças da sílaba “te” e 12 sentenças da sílaba “to”. As demais sentenças destas sílabas apresentam outras métricas, e isso ocorre com muitas das sentenças escandidas, permitindo que uma mesma sentença seja usada para poemas de diferentes métricas, entre hexassílabos até versos bárbaros. A frequência de sentenças por sílabas é exibida parcialmente na Tabela 4.1.

Tabela 4.1: Frequência de Algumas Sentenças por Sílaba.

Sílaba	Qtde.	Sílaba	Qtde.	Sílaba	Qtde.	Sílaba	Qtde.	Sílaba	Qtde.	Sílaba	Qtde.
te.	68	ra.	27	o.	17	ram.	12	To	9	sos.	8
to.	57	vel.	25	ma.	16	ros.	11	os.	9	me.	8
da.	49	va.	21	go.	15	tos.	11	dos	9	gas.	7
do.	45	das.	21	ras.	13	rio.	10	cos.	9	las.	7
se.	32	tes.	20	nha.	13	ga.	10	res.	9	bas.	7
ta.	31	la.	19	tas.	13	ça.	10	O	8	ços.	7
ro.	31	sa.	19	no.	12	rios.	10	#al.	8	po.	6
dos.	31	ç#ão.	18	mo.	12	vam.	10	vo.	8	ro	6
a.	31	co.	17	lo.	12	am.	10	tro.	8	ria.	6
ca.	27	so.	17	de.	12	as.	9	mos.	8	pos.	6

Algumas das sílabas apresentadas na Tabela 4.1, são consideradas rimas agudas, por também serem classificadas como sílabas tônicas, sendo identificadas pelo sinal de cerquilha. Devido ao grande número de sílabas, muitas delas não são apresentadas na Tabela 4.1, mas a Figura 2.10 mostra um número maior destas sílabas, sendo ignoradas as sílabas com apenas um verso, visto que estas não podem ser combinadas para gerar rima.

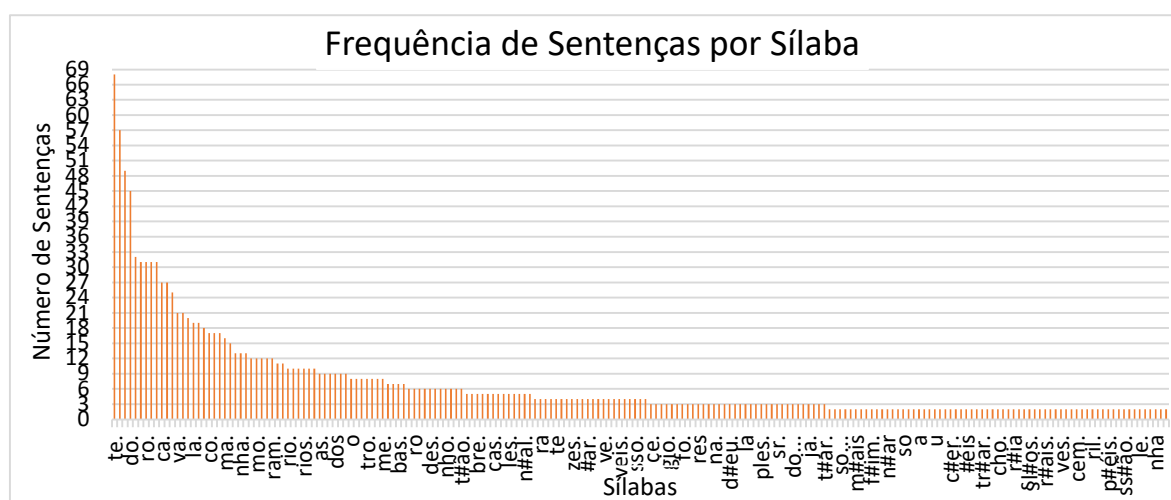


Figura 4.1: Frequência de Algumas Sentenças por Sílabas.

O gráfico da Figura 4.2 apresenta no eixo x as sílabas, e no eixo y a quantidade das sentenças relacionadas às sílabas. Dessa forma, é possível identificar um razoável número de sílabas com baixa frequência de sentenças, e poucas sílabas com disponibilidade para gerar um maior número de combinações. Esta distribuição é um fenômeno empírico, observado apenas com a execução do sistema, e possibilita a tomada de decisão por qual sílaba optar para gerar o poema, além de possibilitar a verificação das possíveis combinações entre as sentenças. O sistema contabilizou a identificação de 203 sílabas consideradas chaves, associadas a 1.428 sentenças, sendo apresentadas no gráfico apenas algumas delas, que também podem ser exibidas por métrica, conforme Figuras 4.2 que exhibe as sentenças decassílabas e Figura 4.3 que mostra as sentenças dodecassílabas.

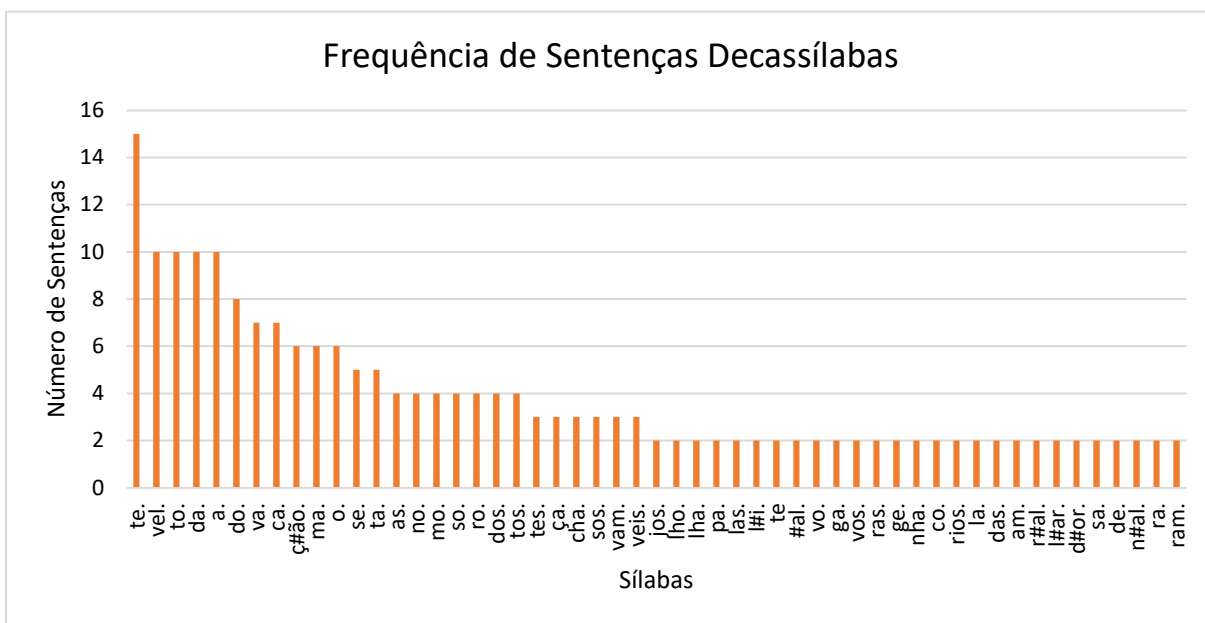


Figura 4.2: Frequência de Sentenças Decassílabas por Sílaba.

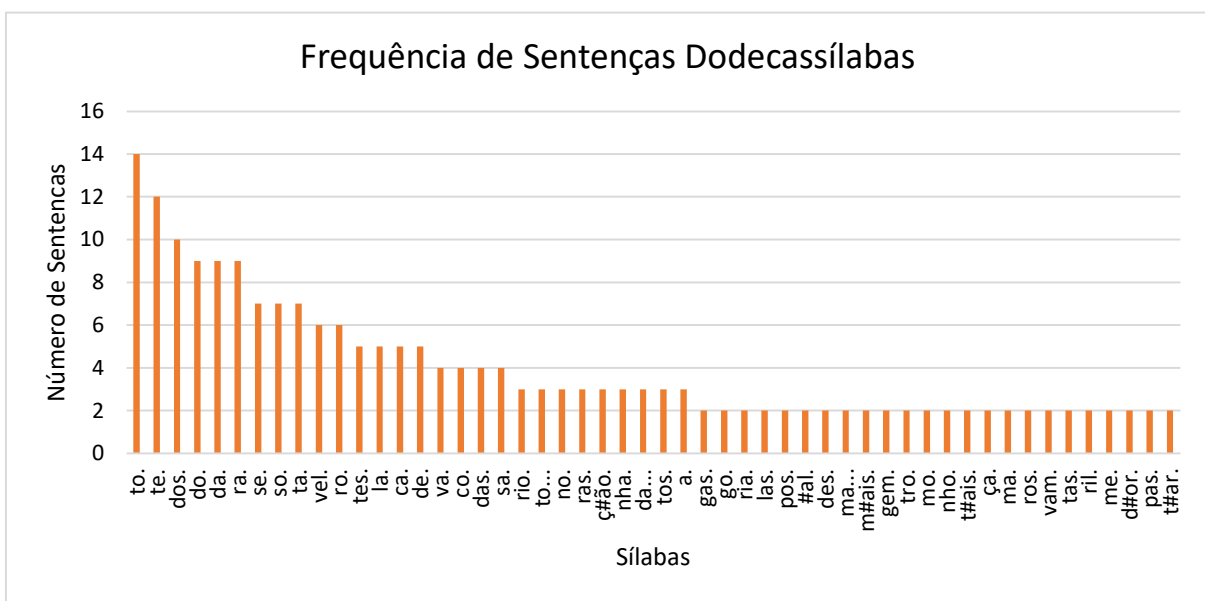


Figura 4.3: Frequência de Sentenças Dodecassílabas por Sílaba.

Nos gráficos acima exibidos, foram ignorados os esquemas rítmicos com apenas uma sentença, por não permitirem rima. Mas, a partir de duas sentenças, os ritmos são exibidos, permitindo perceber em que esquema rítmico as sentenças são mais frequentes e em quais estas são menos frequentes, e é possível realizar muitas possibilidades de combinação de versos com as métricas já mencionadas.

Resultados dos poemas gerados

A Tabela 4.2 apresenta um poema de métrica heptassílaba, como forma de exibir um novo esquema rítmico dos já apresentados até o momento. Este poema foi gerado pelo PROPOE com rima cruzada, formados pela combinação entre sílabas “so” e “vel”, em duas quadras priorizando o esquema rítmico, com referência ao primeiro verso da estrofe. O poema possui um padrão rítmico interessante, de ritmos repetidos, e ainda apresenta coincidência da sílaba “c#ur” em mesma posição de sílaba tônica, sendo as duas palavras agudas, devido ao posicionamento de tais sílabas nas palavras.

Tabela 4.2: Poema de Métrica Heptassílaba.

Poema	Sentenças/Ritmo
Eram o último recurso. Esta é abstrata e irreduzível. É ainda um recurso. Foi um lance admirável.	#E/ram/ O #úl/ti/mo/ re/c#ur/so. – 1 3 7 #Es/ta Ê abs/tr#a/ta E i/rre/du/t#i/vel – 1 3 7 É/ a/#in/da/ um/ re/c#ur/so. – 3 7 F#oi/ um/ l#an/ce/ ad/mi/r#á/vel. – 1 3 7
Foi breve, mas temeroso. A seca é inevitável. Analisemos o caso. Ali era inconcebível.	F#oi/ br#e/ve,/ Mas/ te/me/r#o/so. – 1 2 7 A/ s#e/ca/ Ê i/ne/vi/t#á/vel. – 1 2 A/na/li/s#e/mos/ O/ c#a/so. – 4 7 A/l#i/ #e/ra in/con/cep/t#i/vel. – 2 3 7

Na avaliação feita no poema da Tabela 4.2, a pontuação fonética é de 1,06 pela coincidência entre as letras das palavras que rimam, para acentuação 0,50 considerando, existindo uma tônica similar, e com isso recebe mais 0,33, somado ao esquema rítmico de 0,68, totalizando em 2,57. Considerando que a pontuação máxima para cada poema é de 4 pontos, o resultado alcançado nesse poema corresponde a 64,25%, apresentando uma boa regularidade rítmica, com a presença de todos os critérios estabelecidos para sua qualificação.

Sobre a fonética, são identificadas duas palavras “recurso” coincidindo as letras todas as letras, em “irreduzível” e “admirável”, “inevitável” e “inconcebível” coincidem as letras da rima “vel”, “temeroso” e “caso” coincidem as sílabas que rimam “so”. Quanto à acentuação, todas as últimas palavras de cada verso apresentam rima grave, visto que as sílabas tônicas das palavras que rimam encontram-se na penúltima sílaba.

Conforme o alto número de versos decassílabos na obra de Os Sertões, a Tabela 4.3 apresenta um poema de rima interpolada, com sequência ABBA, em duas estrofes e quatro versos, nas sílabas “ta” e “ma”, cuja similaridade acontece por esquema rítmico, sendo a primeira estrofe marcada por regularidade de posicionamento de sílabas tônicas em 3 7 10 e na segunda estrofe há uma repetição dos posicionamentos 6 e 10. A comparação do esquema rítmico ocorre entre versos, isso significa que cada verso será comparado com o anterior.

Tabela 4.3: Poema de Métrica Decassílaba.

Poema	Sentenças/Ritmo
Quebrara-a de todo a luta. O aspecto reduzia-lhe a fama. A vitória viria por si mesma. Obrigavam-nos a outra luta.	Qu/e/br#a/ra-/a/ de/ t#o/do/ a/ l#u/ta - 3 7 10 O/ as/p#ec/to/ re/du/z#i/a-/l#e a/ f#a/ma.- 3 7 10 A/ vi/t#ó/ri/a/ vi/r#ia/ por/ s#i/ m#es/ma. - 3 7 9 10 O/bri/g#a/vam-/nos/ a/ #o/u/tra/ l#u/ta. - 3 7 10
República a um assomo iconoclasta Esta experiência é belíssima. A princípio avançou corretíssima. Começou a lúgubre visita.	Re/p#ú/bli/ca a um/a/ss#o/mo i/co/no/cl#as/ta- 2 6 10 #Es/ta/ ex/pe/ri/#ên/ci/a É/ be/l#i/ssi/ma. -1 6 10 A/ prin/c#i/pio a/van/ç#o/u/ co/rre/t#i/ssi/ma - 3 6 10 Co/me/ç#o/u/ a/ l#ú/gu/bre/ vi/s#i/ta. - 3 6 10

Para avaliação feita sobre o poema exibido na Tabela 4.3, a soma de fonética é de 0,94 pela coincidência entre as letras das palavras que rimam, acentuação com soma de 0,50 pelo posicionamento das tônicas, com uma tônica similar de pontuação 0,33, esquema rítmico de 0,75, sendo avaliado com pontuação total de 2,48. Nesse poema a soma total corresponde a 62%, apresentando uma boa regularidade rítmica, com a presença de todos os critérios estabelecidos para sua qualificação.

A maior pontuação destes critérios foi o da fonética, visto que no primeiro verso a palavra “luta” se repete, e nas palavras “belíssima” e “certíssima” existem seis letras iguais. Neste caso não houve coincidência entre as sílabas no posicionamento das tônicas e por isso não foi pontuado. Na primeira estrofe todos os versos possuem o mesmo posicionamento de tônicas, sendo que o terceiro verso tem um posicionamento a mais. Já na segunda estrofes, dois dos seus posicionamentos de tônica se repetem em todos os versos.

Uma outra métrica muito evidente na obra de Os Sertões é a dodecassílaba e por isso é apresentado um exemplo na Tabela 4.4. Esse poema foi gerado com base na rima emparelhada, em duas quadras, e a similaridade é feita por posicionamento de tônica, sendo feita a comparação entre os versos.

Tabela 4.4: Poema de Métrica Dodecassílaba.

Poema	Sentenças/Ritmo
Era uma alma belíssima, de valente. Deslumbrava-as ainda o Oriente. Todos saudados; amaldiçoados todos. Mal sobranceava os colmos achatados.	#E/ra/ #u/ma #al/ma/ be/l#i/ssi/ma,/ de/ va/l#en/te. - 1 3 4 7 12 Des/lum/br#a/va-/as/ a/#in/da/ O/ O/ri/#en/te. - 3 7 12 T#o/dos/ s#au/da/dos;/ a/mal/di/ço/#a/dos/ t#o/dos. - 1 3 10 12 M#a/ so/bran/ce/#a/va/ Os/ c#ol/mos/ a/cha/t#a/dos. -1 5 8 12
Ermavam-se as oficinas e as culturas. Insola-se e enregela-se, em 24 horas. As quadras benéficas chegam de improviso. Tomemos um fato, entre muitos, ao acaso.	Er/m#a/vam-/se/ As/ o/fi/c#i/nas/ E As/ cul/t#u/ras. - 2 8 12 In/s#o/la-/se/ E/ en/re/g#e/la-/se, Em/ 24/ h#o/ras. - 2 8 12 As/ qu#a/dras/ be/n#é/fi/cas/ ch#e/gam/ de im/pro/v#i/so.- 2 5 8 12 To/m#e/mos/ um/ f#a/to, #en/tre/ m#u/i/tos,/ ao a/c#a/so.2 5 6 8 12

Na avaliação feita no poema da Tabela 4.4 a soma da fonética é de 0,65 pela coincidência entre as letras das palavras que rimam, para acentuação a soma é de 1,0 por ser um poema composto apenas de rimas graves, para esquema rítmico 0,65, ficando zerado o critério de similaridade de sílabas tônicas e com isso contribuindo para uma redução na pontuação final, totalizando em 2,30. Apesar da ausência de sílabas tônicas similares, esse poema apresenta uma boa regularidade de esquema rítmico com soma total de 57,50%, e isso ocorre devido a pontuação máxima referente ao posicionamento das tônicas nas palavras que rimam.

Nos experimentos realizados na obra de “Os Sertões” foi possível gerar poemas com variadas métricas desde hexassílabos até os versos bárbaros. Assim, considerando que o poema realize

a montagem de quadras, com 296 dodecassílabos, o PROPOE pode gerar 218.556 poemas deste tipo de métrica.

4.2 Recordações do Escrivão Isaías Caminha

A obra “Recordações do Escrivão Isaías Caminha” escrita por Lima Barreto, trata do preconceito racial vivido pelos negros e mulatos após o período da abolição da escravatura, e retrata também a sua insatisfação com a imprensa, pois o seu romance é referenciado com uma auto biografia. Por ser uma obra menor que os Sertões, foi feita a exploração de poucas sentenças, mas com poemas substancialmente relevantes, considerando a rima e o ritmo.

Resultados descritivos dos dados de entrada

Nesta obra foram identificadas 366 sentenças métricas, sendo 60 versos hexassílabos, 40 heptassílabos, 59 Octassílabos, 46 eneassílabos, 44 decassílabos, 48 hendecassílabos, 41 dodecassílabos, 22 bárbaros. No filtro realizado pelo PROPOE, o maior número de sentenças possui terminação “a”, em seguida “da”, e assim sucessivamente até que apresenta muitas sílabas com apenas duas sentenças para representação, o que não garante que tais sentenças sejam de mesma métrica. A Figura 14 apresenta o gráfico da frequência de sentenças por sílaba.

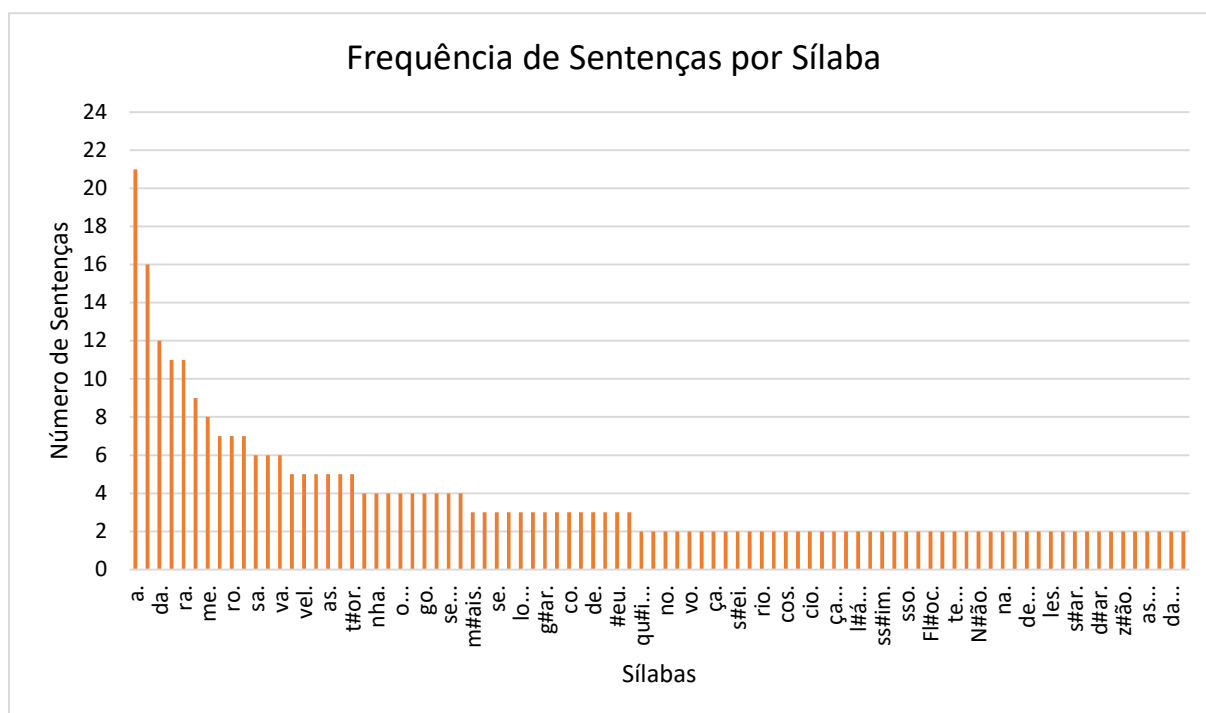


Figura 4.6: Frequência de Sentenças por Sílaba.

As sílabas apresentadas no gráfico, correspondem ao número de sentenças existentes em toda a obra, estas encontram-se escandidas em variadas métricas. Assim, a Figura 4.7 apresenta a frequência de sentenças hexassílabas por métrica, e a Figura 4.8 apresenta as sentenças octossílabas por métrica, identificadas na obra, sendo tais sentenças compostas de uma variedade de sílabas finais a serem combinadas para gerar um poema com rima e ritmo.

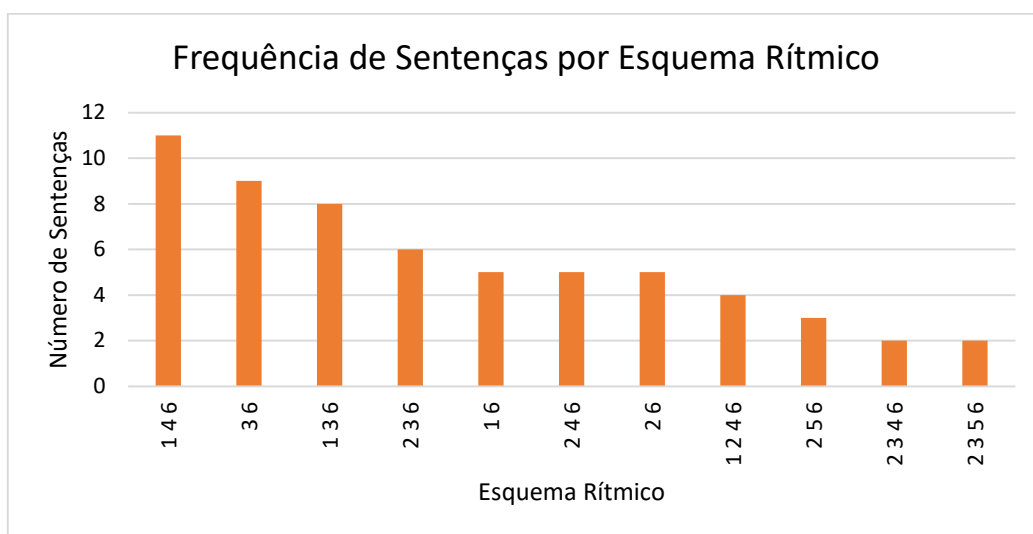


Figura 4.7: Frequência de Sentenças Hexassílabas por Métrica.

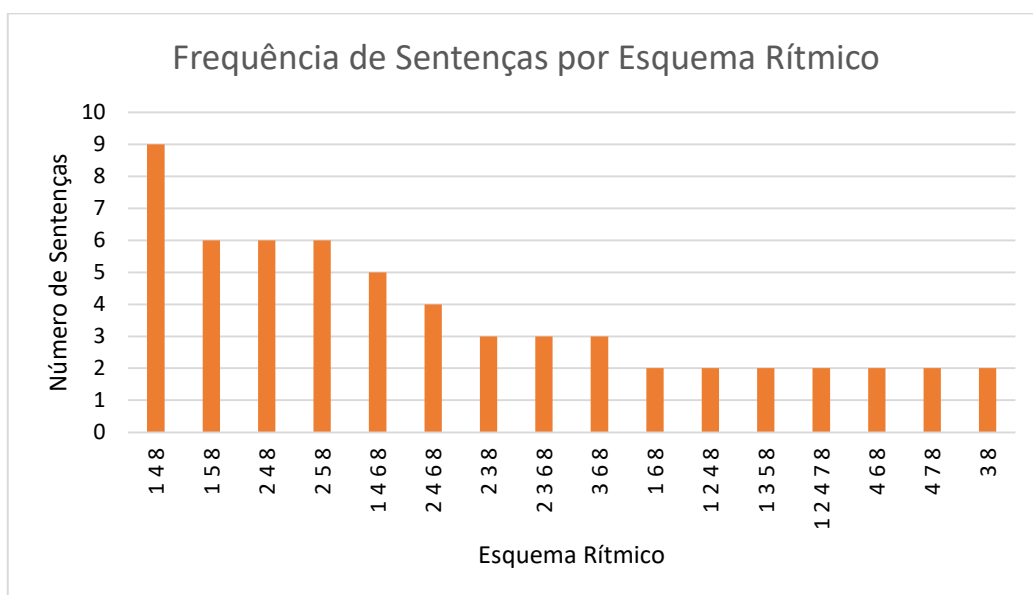


Figura 4.8: Frequência de Sentenças Octossílabas por Métrica.

Resultados dos poemas gerados

Considerando que a maior concentração de sentenças versificadas desta obra é de métricas hexassílabo, a Tabela 4.5 apresenta um poema gerado com esta métrica, sendo uma quadra, com rima de sílabas “z#ão” e “ro”. Para os critérios de montagem, o poema é gerado por esquema rítmico, sendo todos os versos comparados ao primeiro verso da estrofe, e são exibidos em forma de versos e de sentenças para melhor identificação do esquema rítmico e das demais combinações.

Assim, o sistema obtém o filtro das sentenças, faz o cálculo de similaridade entre os versos, sorteia aquele que deve servir de referência aos demais, e segue à procura do próximo verso

que deve ser escolhido pelo esquema rítmico mais similar possível ao primeiro verso, caso este critério contemple mais de um verso, será eleito o que tiver uma acentuação tônica igual.

Tabela 4.5: Poema de Métrica Hexassílabo.

Poema	Ritmo
E ele tinha razão .	E #e/le/ t#i/nha/ ra/z#ão. - 1 3 6
Temo perder dinheiro .	T#e/mo/ per/d#er/ di/nh#e/i/ro.- 1 4 6
E ela tinha razão .	E #e/la/ t#i/nha/ ra/z#ão. - 1 3 6
Bati, veio o porteiro .	Ba/t#i,/ v#e/io O/ por/t#e/i/ro. - 2 3 6

Na etapa de avaliação, o PROPOE identificou duas palavras “razão”, e entre as palavras “dinheiro” e “porteiro” as letras as letras “e i r o”, e o aspecto fonético foi avaliado como 0,90. Sobre o posicionamento das sílabas tônicas, há uma sequência de bom ritmo, pois cada verso possui ao menos dois dos seus posicionamentos iguais ao ritmo do verso comparado, ou seja, o primeiro, sendo que o terceiro verso tem o mesmo esquema rítmico “1 3 6”, com letras iguais, o que gera uma avaliação de 1,67.

Para acentuação de última tônica, foram identificadas duas agudas e duas esdrúxulas, gerando a pontuação de 0,50, contabilizando um total de 3,07 na pontuação do poema, correspondendo a 76,75% do total estabelecido para o valor máximo dos aspectos rítmicos do poema, classificando-o com uma boa pontuação.

Um outro exemplo é apresentado na Tabela 4.6 em duas quadras de métrica bárbara, rima emparelhada, cuja rima é feita com quatro sílabas “da”, “do”, “te” e “a”. A similaridade é por posicionamento de tônica, comparado entre os versos.

Tabela 4.6: Poema de Métrica Bárbara.

Poema	A sua fisionomia estava transtornada. De dia para dia, o jornal crescia em venda. Passei o resto do dia vagueando. O choro a tomou convulsa e eu me afastei chorando. O inspetor repreendeu-as severamente. Apareciam diariamente aos vinte. Demais, a emanção da minha pessoa. A batalha, entretanto, não se decidia.
Ritmo	A/ s#u/a/ fi/si/o/no/m#ia/ es/t#a/va/ trans/tor/n#a/da. 2 8 10 14 De/ d#i/a/ para/ d#i/a, o/ jor/n#al/ cres/c#i/a/ em/ v#en/da. 2 5 8 10 13 Pa/ss#e/i/ o/ r#es/to/ do/ d#i/a/ va/gu/e/#an/do. 2 5 8 13 O/ ch#o/ro a/ to/m#ou/ con/v#ul/sa E #eu/ me a/fas/t#ei/ cho/r#an/do. 2 5 7 8 11 13 O/ ins/pe/t#or/ #re/preen/de/u/as/ se/ve/ra/m#en/te. 4 5 13 A/pa/re/c#i/am/ di/a/ri/a/m#en/te #a/os/ v#in/te. 4 10 12 14 De/m#ais,/ a/ e/ma/na/ç#ã/o/ da/ m#i/nha/ pe/ss#o/a. 2 7 10 13 A/ ba/t#a/lha,/ en/tre/t#an/to,/ n#ão/ se/ de/ci/d#i/a. 3 7 9 13

O poema acima foi avaliado em um total de 2,01, com 0,5 para fonética, sendo identificadas letras iguais em todas as palavras do poema, além das sílabas que rimam. Na acentuação, o poema recebe 1,0 pela presença de tônicas graves em todas as palavras que rimam, e no esquema rítmico 0,51, e com a ausência de sílabas similares nas tônicas, este item fica sem

pontuação. Porém, o poema é classificado com 50,25%, correspondendo a metade do valor estabelecido para o valor máximo dos aspectos rítmicos de um poema.

Devido ao número restrito de sentenças, foram geradas algumas combinações de versos heterométricos, e um deles é exibido na Tabela 4.7, em duas quadras de rima emparelhada, com duas sílabas “a” e “to” a serem rimadas. A similaridade do verso ocorre por esquema rítmicos, sendo o primeiro de cada estrofe tido como referência e os demais seguindo uma regularidade aproximada.

Tabela 4.7: Poema de Versos Heterométricos.

Poema	Ritmo
Podemos ir ao teatro, são oito e meia. Com ele, era assim; não dormia.	Po/d#e/mos/ #ir/ ao/ te/#a/tro,/ São #oi/to E/ m#ei/a. - 2 4 7 9 11 Com/ #e/le,/ #e/ra a/ss#im;/ n#ão/ dor/m#i/a. - 2 4 6 7 9
Loberant não escondia o seu respeito. não serve pra nada, fique certo.	Lo/b#e/rant/ n#ão/ es/con/d#ia/ O/ seu/ res/p#ei/to. - 2 4 7 11 n#ão/ s#er/ve/ pr#a/ n#a/da,/ f#i/que/ c#er/to. - 1 2 4 5 7 9
Era uma mulher andrajosa; parei e ouvi-a. O flanco, que a minha pessoa.	#E/ra #u/ma/ mu/lh#er/ an/dra/j#o/sa;/ pa/r#ei E #o/u/vi/a.- 1 2 5 8 11 O/ fl#an/co,/ que a/ m#i/nha/ pe/ss#o/a. - 2 5 8
Dirigi-me logo para o quarto. Não consegui realizar tudo isto.	D#i/ri/gi/me/ l#o/go/ para O/ qu#ar/to. - 1 5 8 N#ão/ con/se/gu#i/ re/a/li/z#ar/ t#u/do/ #is/to. - 1 4 8 9 11

O poema acima apresenta versos hendecassílabos, eneassílabos e Octassílabos em sua formação. Assim, na avaliação do poema anterior, a fonética foi avaliada basicamente nas sílabas que rimam e o poema recebeu 0,50 para este quesito. Sobre o esquema rítmico o poema recebe 0,61, e sem sílabas iguais nas tônicas dos versos, por isso sem pontuação para sílabas iguais. Quanto a acentuação, apenas uma palavra com acentuação esdrúxula e as demais apresentam acentuação grave, sendo pontuadas com 0,50. O esquema rítmico é bastante variado, mas com sequência interessante em todo o poema devido a posicionamentos iguais em muitos versos, e para este item o poema recebe 0,61, e fica com pontuação total em 1,61, correspondendo ao 40,25% do valor total estabelecido os aspectos rítmicos de um poema. Assim, este poema não apresenta um bom resultado.

4.3 Contrastes e Confrontos

Esta obra é uma coletânea de artigos da imprensa que foram publicados por Euclides da Cunha, que apresentava sugestões para combater a seca no sertão, fazia crítica a degradação ambiental. Além desse tema, o autor também disserta sobre questões trabalhista. Logo, são artigos com abordagens de questões sociais e em defesa dos direitos do trabalhador. Assim, esta obra se difere das demais por se tratar de um novo estilo de texto, e por isso deve ser explorada.

Resultados descritivos dos dados de entrada

A Figura 4.9 apresenta a frequência de sentenças por sílaba identificadas na obra, que apresenta uma variedade pequena de sentenças, mas com potencial para gerar poemas com quatro sílabas, para duas quadra, e neste caso precisa de apenas duas sílabas para cada estrofe.

Nesta obra a maior quantidade de sentenças é de 30 versos heptassílabos, conforme mostra a Figura 4.10, tendo ainda 29 hexassílabos, 28 são octossílabos, 15 são eneassílabos, 14 são

decassílabas, 7 são hendecassílabas e 4 são bárbaras. Na sequência é apresentada na Figura 4.11 a frequência de sentenças hexassílabas por métrica.

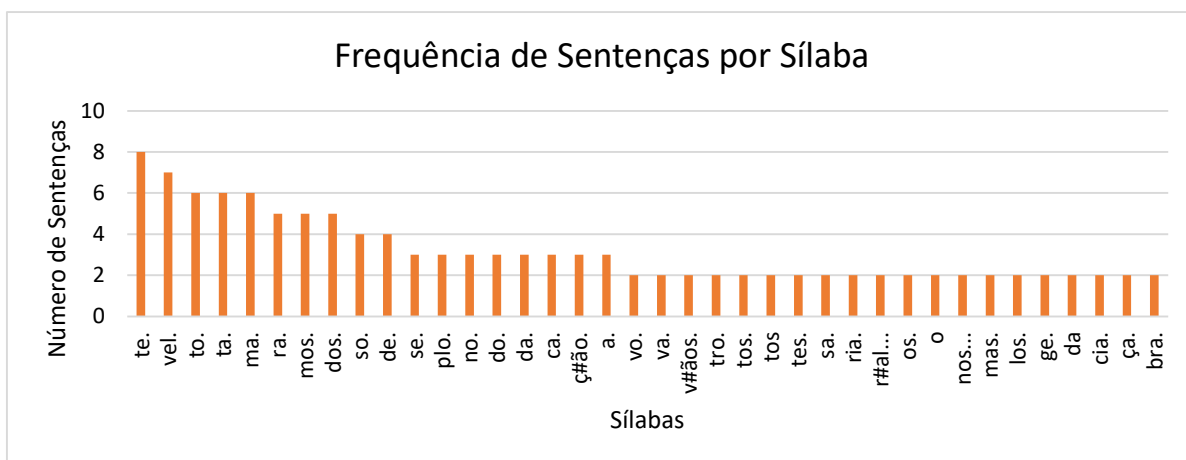


Figura 4.9: Frequência de Sentenças por Sílabas.

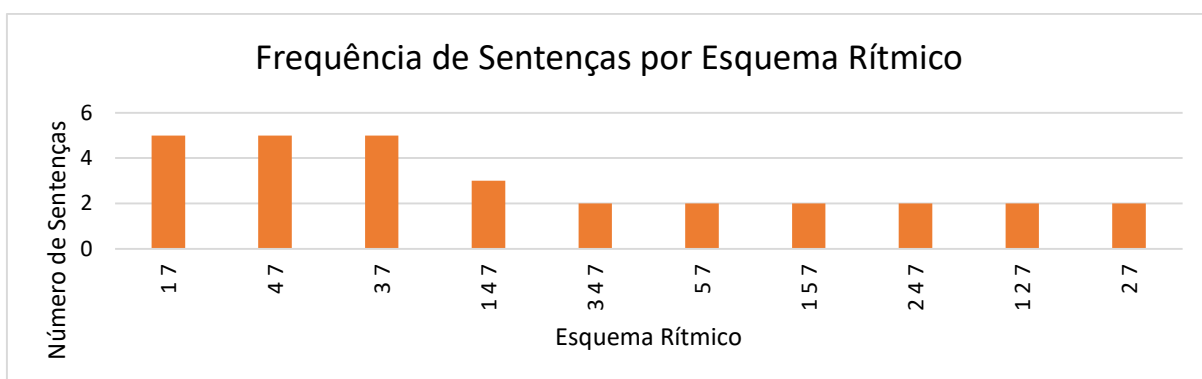


Figura 4.10: Frequência de Sentenças Heptassílabas por Métrica.

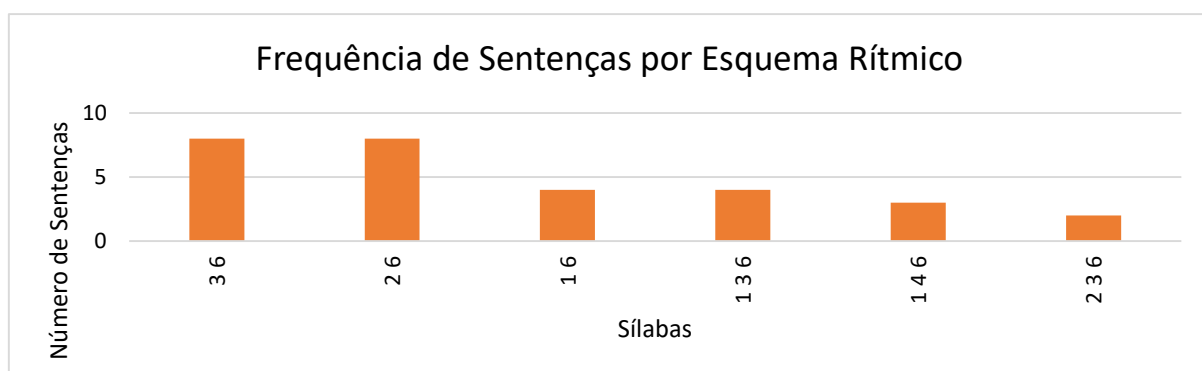


Figura 4.11: Frequência de Sentenças Hexassílabas por Métrica.

Resultados dos poemas gerados

Das sentenças da obra foram feitos experimentos, e um deles é exibido na Tabela 4.8 em forma de poema hexassílabo, em duas quadras, com quatro sílabas a serem rimadas, com similaridade

por esquema rítmico, com versos comparados ao primeiro verso de cada estrofe. Tendo o primeiro verso da primeira estrofe o esquema rítmico 2 6, os demais seguem o padrão aproximado de referência. Como o esquema rítmico da segunda estrofe tem a referência 1 6 um dos versos aproximados coincidiu com o verso de referência da primeira estrofe.

Tabela 4.8: Poema Hexassílabo.

Poema	Ritmo
Contrastes e confrontos	Con/tr#as/tes/ E/ Con/fr#on/tos - 2 6
Ao longo de uma estrada	Ao/ L#on/go/ De #U/ma Es/tr#a/da - 2 4 6
Contrastes e confrontos	Con/tr#as/tes/ E/ Con/fr#on/tos - 2 6
Plano de uma cruzada	Pl#a/no/ De #U/ma/ Cru/z#a/da - 1 3 6
Fronteiam-se e repelem-se.	F#ron/te/i/am/se E/ r#e/pe/lem/se. - 1 6
A Rússia é bárbara.	A/ R#ú/ssi/a/ É/ b#ár/ba/ra. - 2 6
Sublimam-se e acalcanham-se.	S#u/bli/mam/se/ E a/#cal/ca/nham/se. - 1 6
A situação é clara.	A/ si/tua/ç#ã/o É/ cl#a/ra. - 4 6

O poema anterior foi avaliado alcançando um total de 3,10. Para esquema rítmico, o poema foi avaliado com 0,60, e com 1,0 ponto pela igualdade nas sílabas tônicas de “tras” e “fron”. Para acentuação o poema tem 3 pontos, dividido por 4 combinações resultando em 0,75. A fonética tem 1,0, devido a quantidade de letras repetidas nas palavras que rimam, totalizando uma pontuação que corresponde a 77,50% do valor total estabelecido os aspectos rítmicos de um poema, sendo este considerado um bom resultado.

Outro poema é apresentado na Tabela 4.9, em forma de duas quadras de rima mista de sequência ABBA e CDCD, em métrica heptassílabo, com quatro sílabas para rimar. O verso é montado pelo critério de esquema rítmico, com a comparação dos versos com base no primeiro verso de cada estrofe. Nesse caso, os versos do poema são comparados ao primeiro verso de cada estrofe gerando uma pontuação de 0,54. Além disso, o poema recebe 0,75 pela acentuação de muitas tônicas graves, e 0,65 de fonética com letras iguais em palavras que rimam, totalizando 1,94 de pontuação para o poema, sendo classificado com 48,50% do valor total estabelecido os aspectos rítmicos de um poema, não sendo considerado como um bom resultado.

Tabela 4.9: Poema Heptassílabo.

Poema	Ritmo
É uma sugestão empolgante.	É #u/ma/ su/ges/t#ão em/pol/g#an/te. - 1 5 7
Lança o olhar em volta.	L#an/ça/ O/ o/lh#ar/ Em/ v#ol/ta. - 1 5 7
Uma obsessão de artista.	#U/ma ob/se/ss#ã/o/ de ar/t#is/ta. - 1 4 7
É um contraste comovente.	É um/ con/tr#as/te/ co/mo/v#en/te. - 3 7
Era um recuo deplorável.	#E/ra um/ re/c#uo/ de/plo/r#á/vel. - 1 4 7
Foi um triunfo e um revide.	F#oi/ um/ tri/#un/fo E um/ re/v#i/de. - 1 4 7
Sobretudo explicável.	So/bre/t#u/do/ ex/pli/c#á/vel. - 3 7
O ianque aplaude.	O/ i/#an/qu/e/ a/pl#a/u/de. - 3 7

O poema heterométrico é apresentado na Tabela 4.10, em duas quadras de rima emparelhada, e quatro sílabas a serem rimadas no poema. A sua similaridade é por esquema rítmico, com

comparação dos versos com o primeiro verso de cada estrofe. Por apresentar uma variedade métrica, poucas são as coincidências entre os esquemas rítmicos.

Tabela 4.10: Poema Heterométrico.

Poema	Aceitava-se um a um; repelia-nos unidos. As secas do Norte interessam a dez Estados. É da própria ordem física. Trata-se de uma pantomima heróica. Está numa alternativa. Esta última consideração é expressiva. Fronteiam-se e repelem-se. Sublimam-se e acalcanham-se.
Ritmo	A/#ce/i/ta/va/se um/ a um;/ r#e/pe/lia/nos/ u/n#i/dos. - 2 8 13 As/ s#e/cas/ do/ N#or/te in/te/r#e/ssam/ a/ d#ez/ Es/t#a/dos. - 2 5 8 11 13 É/ da/ pr#ó/pri/a/ #or/dem/ f#i/si/ca. - 3 6 8 Tr#a/ta-/se/ de #u/ma/ pan/to/m#i/ma he/r#ó/ca. - 1 4 8 10 Es/t#á/ n#u/ma al/ter/na/t#i/va. - 2 3 7 #Es/ta #úl/ti/ma/ con/si/de/ra/ç#ã/o É ex/pre/ss#i/va. - 1 2 9 12 F#ron/te/i/am/se/ E/ r#e/pe/lem/se. - 1 7 S#u/bli/mam/se/ E/ a/#cal/ca/nham/se. - 1 7

O poema é avaliado com 0,47 para esquema rítmico, 0,20 para sílabas iguais em posicionamento de tônica, 0,75 para acentuação, composta por apenas duas esdrúxulas e as demais tônicas graves. A fonética recebe pontuação de 0,58, tendo algumas das suas letras iguais em palavras, além das sílabas que rimam. Recebendo um total de 50% em sua pontuação, este percentual correspondendo a metade do valor total estabelecido para os aspectos rítmicos de um poema, sendo este considerado um bom resultado.

4.4 Os Lusíadas

Os Lusíadas é uma obra de poesia épica do escritor português Luís Vaz de Camões que narra as perigosas trajetórias marítimas de Vasco da Gama às Índias, bem como a descoberta de novas terras. O seu livro está dividido em dez contos organizados em 1.102 estrofes, cada uma com oito versos decassílabos. Esta obra foi considerada uma arte devido ao empenho do autor em manter uma regularidade nos seus versos, sendo o maior poema épico da língua portuguesa.

Resultados descritivos dos dados de entrada

Dessa obra o PROPOE recebe como entrada um alto número de versos, com muitas possibilidades de combinações, cujas maiores frequências de sentenças por sílaba são de métrica decassílaba com 3.019 sentenças, 2.212 sentenças hendecassílabas, 1.098 sentenças dodecassílabas, e 226 sentenças bárbaras.

A Figura 4.12 mostra a frequência de sentenças por sílaba através de um gráfico, com 424 sílabas “do”, sendo exibido uma grande quantidade de sentenças de uma mesma sílaba, e poucas sentenças para um grande número de sílabas, ou seja, grandezas inversamente proporcionais.

O sistema também recebe um número significativo de sentenças de métrica hendecassílabas, com grande número de sentenças correspondendo a poucos esquemas rítmicos e um pequeno número de sentenças em grande variedade de esquemas rítmicos conforme Figura 4.14.

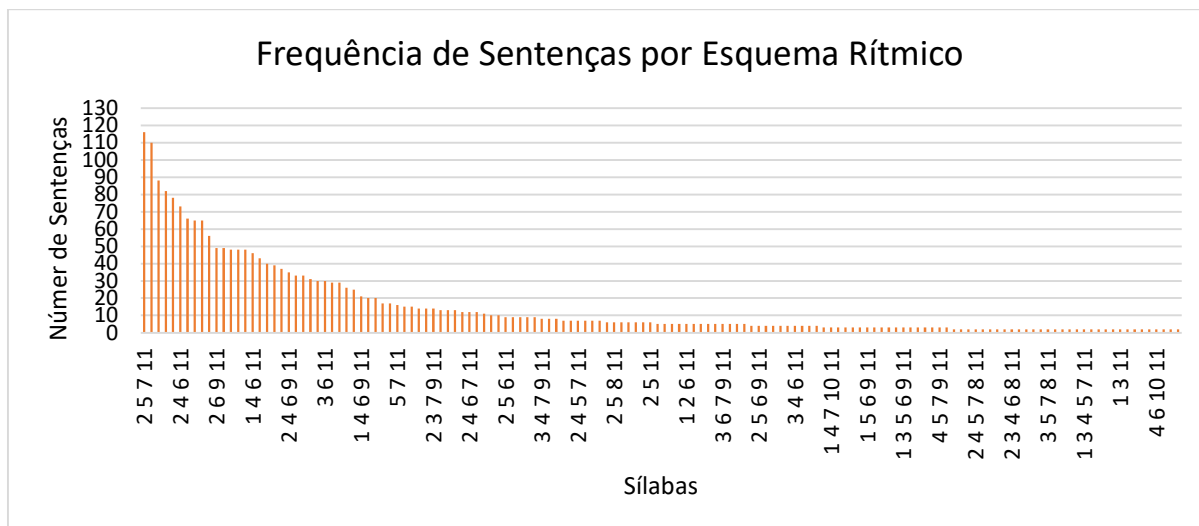


Figura 4.14: Frequência de Sentenças Hendecassílabas por Métrica.

Resultados dos poemas gerados

A Tabela 4.11 mostra um dos poemas gerados a partir da obra de Os Lusíadas, em duas quadras de versos decassílabos, com rima cruzada, de sílabas “do” e “te”, cuja similaridade de tônica é com base no seu posicionamento, sendo comparados entre versos. Porém, o primeiro verso de cada estrofe se mantém como referência para o próximo verso candidato.

Tabela 4.11: Poema Decassílabo.

Poema	Ritmo
Tem pera vós por dote aparelhado. Vós, tenro e novo ramo florecente. Já a Ilha, e todo o mais, deseparando. Vêm pela Via Láctea juntamente.	T#em/ p#e/ra/ v#ós/ por/ d#o/te a/pa/re/lh#a/do. - 1 2 4 6 10 V#ós,/ t#em/ro E/ n#o/vo/ r#a/mo/ flo/re/c#en/te. - 1 2 4 6 10 J#á/ a #I/lha, E/ t#o/do O/ m#ais,/ de/sem/pa/r#an/do. - 1 2 4 6 10 V#êm/ p#e/la/ V#i/a/ L#ác/tea/ jun/ta/m#en/te. - 1 2 4 6 10
Dai-me agora um som alto e sublimado. De outra pedra mais clara que diamante. Pera ir buscar o Indo desejado. Que outra Ilha tem perto, cuja gente.	D#ai/me a/g#o/ra um/ s#om/ #al/to E/ su/bli/m#a/do. - 1 3 5 6 10 De #ou/tra/ p#e/dra/ m#ais/ cl#a/ra/ que/ dia/m#an/te. - 1 3 5 6 10 P#e/ra/ #ir/ bus/c#ar/ O #In/do/ de/se/j#a/do. - 1 3 5 6 10 Que #o/u/tra #I/lha/ t#em/ p#er/to,/ c#u/ja/ g#en/te. - 1 3 5 6 8 10

No poema acima foi identificado um esquema rítmico que segue igual para todos os versos da primeira estrofe, e para os versos da segunda estrofe, muda apenas um dos posicionamentos da tônica no último verso. Por isso, a pontuação para este critério é de 0,97. Quanto à fonética algumas das letras além das que rimam são iguais entre as palavras, e isso gera uma pontuação de 0,60. Todas as últimas palavras dos versos possuem acentuação de sílaba grave, e com isso a o poema recebe mais 1,0 ponto, totalizando esta pontuação em 2,57, que corresponde a 64,25% do total estabelecido para os aspectos rítmicos de um poema, sendo este considerado um bom resultado.

Um outro exemplo de poema desta obra é apresentado na Tabela 4.12, em duas quadras de versos hendecassílabos, de similaridade por posicionamento de tônica, com comparação entre versos, sendo o primeiro verso de cada estrofe servindo de referência para o verso candidato.

Tabela 4.12: Poema Hendecassílabo.

Poema	Ritmo
Do Rei, que já sabia a gente que era. Nas águas têm passado o duro Inverno. Sagaz, astuto e sábio em todo o dano. Não entra pela barra, e surge fora.	Do/ R#ei./ que/ j#á/ sa/b#i/a/ a/ g#en/te/ que #e/ra. - 2 4 6 9 11 Nas/ #á/guas/ t#êm/ pa/ss#a/do/ O/ d#u/ro In/v#er/no. - 2 4 6 9 11 Sa/g#az./ as/t#u/to E/ s#á/bi/o Em/ t#o/do O/ d#a/no. - 2 4 6 9 11 N#ão/ #en/tra/ p#e/la/ b#a/rra./ E/ s#ur/ge/ f#o/ra. - 1 2 4 6 9 11
Que o malévolo Baco lhe ensinara. E sem a fúria horrenda de Vulcano. Por um pregão do ninho meu paterno. A pertinácia aqui lhe custa cara.	Que/ O/ ma/l#é/vo/lo/ B#a/co/ lhe en/si/n#a/ra. - 4 7 11 E/ Sem/ a/ f#ú/ri/a ho/rr#en/da/ de/ Vul/c#a/no. - 4 7 11 Por/ um/ pre/g#ã/o/ do/ n#i/nho/ m#eu/ pa/t#er/no. - 4 7 9 11 A/ per/ti/n#á/ci/a a/qu#i/ lhe/ c#us/ta/ c#a/ra. - 4 7 9 11

O poema da Tabela 4,12 evidencia bem o sorteio do primeiro verso de cada estrofe, devido ao esquema rítmico de cada um desses versos. Na primeira estrofe os ritmos seguem basicamente o mesmo esquema do primeiro verso, mesmo sendo comparado o anterior com o posterior. Já na segunda estrofe acontecem mudanças nos esquemas rítmicos, apenas com acréscimo de uma tônica, sendo avaliado com 0,93. Na fonética as letras pontuadas são basicamente as que rimam, e recebem 0,75 na pontuação. A acentuação recebe 1,0 ponto, pois o poema é formado por tônicas graves, totalizando em 2,68 pontos correspondendo a 67% do valor total estabelecido para os aspectos rítmicos de um poema, sendo este considerado um bom resultado.

Na Tabela 4.13 é apresentado um poema dodecassílabo, formado por duas quadras em rima cruzada, com sílabas “ram” e “tes”, com similaridade por posicionamento de tônica, comparados ao primeiro verso de cada estrofe.

Tabela 4.13: Poema Dodecassílabo.

Poema	Ritmo
Que os Mouros cautelosos se guardaram. Tentou Perito e Teseu, de ignorantes. Do Reino onde as Hespéridas viveram. Em prêmio destes feitos excelentes.	Que Os/ M#o/u/ros/ c#a/u/te/lo/sos/ se/ guar/d#a/ram. - 2 5 12 Ten/t#o/u/ Pe/r#i/to/ E/ Te/s#eu./ de ig/no/r#an/tes. - 2 5 9 12 Do/ R#e/i/no/ #on/de As/ Hes/p#é/ri/das/ vi/v#e/ram. - 2 5 8 12 Em/ pr#é/mi/o/ d#es/tes/ f#e/i/tos/ e/xce/l#en/tes. - 2 5 7 12
E, enfim, co Bétis tanto alguns puderam. Viu de antigos, longincos e altos montes. Que à terra, de Vandália nome deram. Espedaçam-se as lanças, e as frequentes.	E./ en/f#im./ c#o/ B#é/tis/ t#an/to/ al/g#uns/ pu/d#e/ram3 4 5 7 10 12 V#iu/ de/ an/t#i/gos./ lon/g#in/cos/ E/ #al/tos/ m#on/tes. - 1 4 7 10 12 Que/ à/ t#e/rra./ de/ Van/d#á/li/a/ n#o/me/ d#e/ram. - 3 7 10 12 Es/pe/d#a/çam/-se/ As/ l#an/ças./ E/ As/ fre/qu#en/tes. - 3 7 12

O poema da Tabela 4.13 apresenta pontuação para esquema rítmico de 0,66, para fonética 0,80, para acentuação 1,0, pois todas são tônicas graves. Para similaridades de tônicas, foi encontrada a sílaba “da” no mesmo posicionamento de tônica do primeiro e do terceiro versos da segunda estrofe, e por isso esse poema recebe 0,17 pontos, totalizando em 2,13 pontos para todo o poema correspondendo a 53,25% do valor total estabelecido para os aspectos rítmicos de um poema, sendo este considerado um bom resultado.

Capítulo 5

Considerações Finais

“A maior recompense para o trabalho do homem não é o que ele ganha com isso, mas o que ele se torna com isso.”

-- John Ruskin

Neste trabalho, foram realizadas pesquisas ao longo da sua construção, que permitiram o cumprimento do objetivo proposto, sobre o desenvolvimento de um sistema computacional capaz de organizar sentenças metrificadas de estruturas de versificação da prosa literária brasileira para gerar poemas, com base em critérios e padrões rítmicos estabelecidos para esta finalidade. Com isso, foi implementado o PROPOE, considerada uma ferramenta para gerar poemas computacionais, tendo como entrada as sentenças versificadas de obras literárias brasileiras, que são filtradas para reduzir o número de sentenças versificadas, a serem combinadas na geração do poema, passando por uma avaliação feita pelo próprio sistema, relacionado a aspectos rítmicos.

No contexto dessa pesquisa existe um elo entre as áreas da tecnologia e da produção humana e por isso considerado como uma atividade de humanidade digital. Além da obra Os Sertões, o sistema recebeu dados das prosas literárias brasileiras de Recordações do escrivão Isaías Caminha, Contrastes e Confrontos, e dessas obras foram feitas variadas combinações e identificados esquemas com boa correspondência entre ritmos dos versos gerados. Foi feito também o experimento com a obra literária Os Lusíadas - composto por versos metrificados -, de onde o PROPOE realizou recombinações dos seus versos, gerando novos poemas, mas sempre considerando as ideias do poeta.

O processo de gerar tais poemas automaticamente agiliza uma atividade que pode ser desempenhada também através de procedimentos não-computacionais, pois o sistema realiza correlações entre um grande número de sentenças de cada obra literária, esgotando as possibilidades de combinações que podem passar despercebidas por humanos. Dessa forma o sistema tem como referência a *Augmented Intelligence*, por realizar atividade exaustiva de combinações, oferecendo o poder de decisão ao humano na escolha do poema.

Os poemas gerados foram avaliados pelo próprio sistema, com base na coincidência entre as letras das palavras que rimam, bem como as suas sílabas tônicas, sobre a coincidência do posicionamento da sílaba tônicas, e sobre a coincidência do esquema rítmico entre pares de versos. Assim, para cada um destes critérios o poema recebe um ponto por coincidência identificada, mas em seguida as pontuações são normalizadas para que cada poema tenha no máximo a pontuação 4, ou seja, 1 ponto por critério.

Sobre os critérios considerados para avaliação, todos os poemas foram pontuados em pelo menos três deles, ocorrendo em alguns poemas a ausência na coincidência entre as sílabas tônicas dos esquemas rítmicos, mas alcançando pontuações consideradas aceitáveis para poemas com um bom ritmo, devido a soma dos demais critérios. Além disso, foi possível realizar um grande número de variadas combinações entre sentenças, e isso significa que os poemas atendem em algum nível às regras estabelecidas pela gramática para geração dos mesmos. Apesar de não ter feito a verificação da semântica, os poemas podem fazer sentido. Tal aspecto é ignorado no sistema por se partir do pressuposto de que os versos serão combinados sempre de determinada obra literária, quando as palavras se correlacionam em um mesmo contexto.

Quando comparado aos trabalhos anteriores, o PROPOE é o primeiro sistema para o português brasileiro. Além disso, é o primeiro sistema a gerar poemas a partir de sentenças versificadas de prosa literária. A montagem do poema realiza a otimização combinatória de similaridade destes versos por meio de um algoritmo guloso. Tal estratégia permite a redução do custo computacional, por permitir a comparação apenas das sentenças candidatas, gerando poemas com variados tipos de rima, de forma isométrica ou heterométrica, possibilitando escolha de duas ou quatro sílabas, e sendo avaliando pelos aspectos relevantes que caracterizam um poema regular. Além disso, o PROPOE é o único sistema a realizar uma maior exploração de aspectos rítmicos em poemas, principalmente na comparação das letras entre as palavras que rimam, bem como no posicionamento da sílaba tônica quanto a acentuação.

Sobre as limitações do sistema, é possível destacar a sua dependência do sistema MIVES que exporta os dados extraídos de obras literárias brasileiras, visto que o PROPOE recebe estes dados como entrada. Além disso, os demais aspectos que limitam o sistema correspondem a ausência de avaliação feita por humanos, por reconhecer que estes podem contribuir com melhorias no sistema. Outro fator limitante corresponde ao pequeno número de obras literárias exploradas que serviram como experimento para alcance dos resultados.

Apesar de considerar como pequeno, o número de obras exploradas pelo sistema, é possível explorar muitas outras, desde que tenham sido versificadas pelo sistema MIVES. Como trabalhos futuros existe a proposta de ampliar o número de exploração de obras literárias, fazer identificação de aspectos relacionados a rima interna, atribuir aspectos de sentimento, realizar avaliação dos poemas por diferentes grupos de pessoas, e avaliar a semântica dos poemas gerados.

Os dados extraídos do PROPOE servem como análise para pesquisas futuras, por este apresentar estruturas poéticas de variadas obras literárias e permitir a combinação dos segmentos frásicos. No anexo 1 são apresentados alguns dos poemas gerados pelo PROPOE, com um breve resumo relacionado a avaliação de cada um deles, e a apresentação tanto na sua forma estética, quanto na forma de sentença metrificada para melhor verificação da métrica e do seu esquema rítmico.

Referências Bibliográficas

[Almeida 1946] Almeida G. de, (1946). Guardados da Memória - A poesia de “Os sertões”. p. 205-228.

[Antonio 2016] Antonio J. L., (2016). Produções Poéticas com o(s) computador(es). SIBILA: Revista de poesia e crítica literária. Ano 16. ISSN 1860-289X. 2009.

[Agirrezabal 2015] Agirrezabal, M. et. al., (2013). POS-tag based poetry generation with wordnet. In Proceedings of the 14th European Workshop on Natural Language Generation, pages 162–166, Sofia, Bulgaria, August. ACL Press.

[Boden 2009] Boden, A. M. (2009). Association for the Advancement of Artificial Intelligence. All rights reserved. ISSN 0738-4602.

[Campos 1996] Campos A. de., (1996). Transertões. Folha de São Paulo. Citado 6 vezes nas páginas 5, 7, 9, 21, 22 e 36./

[Carvalho, Loula e Queiroz 2016] Carvalho, R., Loula, A. e Queiroz, J., (2016). Poetry in Prose: automatic identification of verses in brazilian literature. Disponível em: <<http://dh2016.adho.org/abstracts/290>>.

[Chrismartin, Tobing, e Manurung 2015] Chrismartin, B.; Tobing, L.; And Manurung, R., (2015). A chart generation system for topical metrical poetry. In Proceedings of the 6th International Conference on Computational Creativity, Park City, Utah, USA.

[Cunha e Cintra 2017] Cunha, C.; Cintra, L., (2017). Nova gramática do português contemporâneo. Rio de Janeiro: Lexikon Editora Digital, p. 685-727.

[Das e Gambäck 2014] Das, A.; Gambäck, B. (2014). Poetic Machine: Computational Creativity for Automatic Poetry Generation in Bengali.

[D’Onofrio 2016] D’Onofrio, S., (2016). Forma e sentido do texto literário. São Paulo: Ática, 2007. p. 180-271. GHAZVININEJAD, M., et. al. Generating Topical Poetry. Conference on Epirical Methods in Natural Language Processing, p. 1183-1191, Austin, Texas.

[Ghazvininejad et. al., 2016] Ghazvininejad *et. al.*, (2016). Generating Topical Poetry. Conference Methods on Natural Language Processing.

[Gervás 2000] Gervás, P., (2000). WASP: Evaluation of Different Strategies for the Automatic Generation of Spanish Verse. Proceedings of the AISB-00 symposium on creative & cultural aspects of AI. p. 93-100.

- [Gervás 2001] Gervás, P., (2001). An Expert System for the Composition of Formal Spanish Poetry. *Journal of Knowledge-Based Systems*. v. 14, Ed. Elsevier. p 181-188.
- [Gervás *et. al.* 2002] Gervás, P. *et. al.* (2002). Poetry Generation in Colibri. *European Conference on Case-Base Reasoning*. p. 73-87.
- [Gervás 2013] Gervás, P. (2013). Computational Modelling of Poetry Generation. *Conference Artificial Intelligence and Poetry Symposium, AISB Convention*.
- [Gervás 2016] Gervás, P. (2016). Constrained creation of poetic forms during theme-driven exploration of a domain defined by an N-gram model. *Journal Connection Science*, p. 111-130.
- [Goldstein 2006] Goldstein, N. S., (2006). *Versos, sons, ritmos*. v. 14. ed. São Paulo: Ática, p. 64 – 98.
- [Lutz 1959] Lutz, Theo., (1959). *Stochastische Texte*. Disponível em: <http://www.stuttgarter-schule.de/lutz_schule_en.htm>.
- [Manurung, Ritchie e Thompson 2000] Manurung, H. M.; Ritchie, G.; Thompson, H., (2000). *Towards A Computational Model Of Poetry Generation*. The University of Edinburgh.
- [Manurung 2004] Manurung, H. (2004). An evolutionary algoritmo approach to poetry generation. University of Edinburgh. College of Science and Engineering. School of Informatics.
- [Misztal e Indurkha 2014] Misztal J. e Indurkha B. (2014) Poetry generation system with an emotional personality. In *5th International Conference on Computational Creativity, 2014*, Ljubljana, Slovenia.
- [Oliveira 2009] Oliveira, H. G. (2009). *Automatic Generation of Poetry: an overview*. Universidade de Coimbra, Portugal. CISUC.
- [Oliveira 2012] Oliveira, H. G.(2012). *PoeTryMe: a versatile platform form poetry generation*. “Computational Criativity, Concept Invention and General Intelligence”.
- [Oliveira 2015] Oliveira, H. G. (2015). *Tra-la-Lyrics 2.0: Automatic Generation of Song Lyrics on a Semantic Domain*. *Journal of Artificial General Intelligence*, 2015.
- [Oliveira 2017] Oliveira, H. G. A. (2017). *Survey on Intelligent Poetry Generation: Languages, Features, Techniques, Reutilisation and Evaluation*. In *10h International Conference on Natural Language Generation, 2017*.
- [Oliveira *et. al.* 2017] Oliveira, H. G. A. *et. al.* (2017). *Multilingual Extension and Evaluation of a Poetry Generator*. *Journal Natural Language Engineering*. p. 919-967.
- [Proença 2007] Proença, F. D. (2007). *A Linguagem Literária*. 8. ed. – São Paulo: Ática, p. 62-79.

[Queneau 1961] Queneau, R. (1961). 100.000.000.000.000 de poemes. Gallimard Series. Schoenhof's Foreign Books, Incorporated.

[Rashel e Manurung 2014] Rashel, F.; Manurung, R. (2014). Pemuisi: a constraint satisfaction-based generator of topical Indonesian poetry. In Proceedings of 5th International Conference on Computational Creativity, Ljubljana, Slovenia. jun, 2014. p. 82-90.

[Rojas 1996] R. A Systematic (1996). Neural Networks: A Systematic Introduction. MathWorks: Documentation. Disponível em: <<https://www.mathworks.com/help/nnet/gs/fit-data-with-a-neural-network.html>>. Acesso em maio 2018.

[Silva 2014] Silva, P. P.(2014). Teoria da Literatura I. Biblioteca Universitária Pearson. São Paulo: Pearson Education do Brasil, p. 96 – 127.

[Yan 2016] Yan, R. (2016). I, Poet: Automatic poetry composition through recurrent neural networks with iterative polishing schema. In Proceedings of 25th International Joint Conference on Artificial Intelligence, IJCAI 2016, pages 2238-2244. New York, NY, USA, July. IJCAI/AAAI Press.

[Zhang e Lapata 2016] Zhang, X.; Lapata, M. (2016). Chinese poetry generation with recurrent neural network. In Proceedings of 2014 Conference on Empirical Methods on Natural Language Processing. Pages 670-680. Doha, Qatar, October. ACL Press.

[Jiang e Zhou 2008] Jiang, L.; Zhou, M. (2008). Generating Chinese Couplets using a Statistical MT Approach. Microsoft Reserch Asia.

Anexo 01

Poemas gerados durante os experimentos

Os Poemas apresentados a seguir foram gerados pelo PROPOE, e mostram as combinações realizadas entre os versos de cada obra literária. Os poemas exibidos em forma de quadras, extradidos dos livros Os Sertões, Recordações do escrivão Isaías Caminha, Contrastes e Confrontos e Os Lusíadas. Os poemas são exibidos também em forma escandida para melhor identificação da separação em sílabas poéticas e verificação do posicionamento das sílabas tônicas.

Os Sertoes

Poema 1

O poema 1 é de rima cruzada, métrica decassílaba, duas estrofes e quatro versos, com sílabas “sos” e “do” a serem rimadas, similaridade por esquema rítmico, comparação de versos com o primeiro verso de cada estrofe. Como avaliação o poema recebe 0,50 de fonética, 0,0 de tônicas similares, 0,46 de esquema rítmico e 0,50 de acentuação, gerando um total de 1,46 que corresponde a 37% do total geral, classificando o poema com baixa pontuação,

Esquadrinhava os jiraus suspensos.
O fato foi de todo inesperado.
Vozeavam aclamações e aplausos.
Reagiu apesar de ferido.

Destruíra e criara revoltosos.
Era um golpe de audácia estupendo.
Não eram sitiantes, eram presos.
Não havia errar o alvo desmedido.

Es/qua/dri/nh#a/va/ Os/ ji/r#aus/ sus/p#en/sos. -> Decassílabo -> 4 8 10

O/ f#a/to/ f#oi/ de/ t#o/do i/nes/pe/r#a/do. -> Decassílabo -> 2 4 6 10

Vo/ze/#a/vam/ a/c/a/ma/ç#ões/ E a/pl#au/sos. -> Decassílabo -> 3 8 10

Re/a/g#i/u/ a/pe/s#ar/ de/ fe/r#i/do. -> Decassílabo -> 3 7 10

Des/tru/#i/ra E/ cri/#a/ra/ re/vol/t#o/sos. -> Decassílabo -> 3 6 10

#E/ra um/ g#ol/pe/ de au/d#á/ci/a es/tu/p#en/do. -> Decassílabo -> 1 3 6 10

N#ão/ #e/ram/ si/ti/#an/tes,/ #e/ram/ pr#e/sos. -> Decassílabo -> 1 2 6 8 10

N#ão/ ha/v#ia/ e/rr#ar/ O #al/vo/ des/me/d#i/do. -> Decassílabo -> 1 3 5 6 10

Anexo

Poema 2

O poema 2 é de rima cruzada, métrica decassílaba, duas estrofes e quatro versos, com sílabas “vel” e “dos” a serem rimadas, similaridade por esquema rítmico, comparação de versos com o primeiro verso de cada estrofe. Como avaliação o poema recebe 0,83 de fonética, 0,33 de tônicas similares, 0,62 de esquema rítmico e 0,50 de acentuação, gerando um total de 2,28, que corresponde a 57% do valor geral, classificando o poema com boa pontuação.

Esta impõe-se, tenaz e inflexível.

Tinha um plano único — ir a Canudos.

Não havia errar o alvo imóvel.

Perdiam-se nos casebres unidos.

Previram a derrota inevitável.

À noite tiroteios cerrados.

O antagonismo era inevitável.

Os assaltantes eram repelidos.

#Es/ta im/p#ô/e/§se,/ te/n#az/ E in/fle/x#i/vel. -> Decassílabo -> 1 3 7 10

T#i/nha/ um/ pl#a/no #ú/ni/co — #ir/ a/ Ca/n#u/dos. -> Decassílabo -> 1 4 5 7 10

N#ão/ ha/v#ia/ e/rr#ar/ O/ #al/vo/ i/m#ó/vel. -> Decassílabo -> 1 3 5 7 10

Per/d#i/am§/se/ nos/ ca/s#e/bres/ u/n#i/dos. -> Decassílabo -> 2 7 10

Pre/v#i/ram/ a/ de/rr#o/ta i/ne/vi/t#á/vel. -> Decassílabo -> 2 6 10

À/ n#oi/te/ ti/ro/t#e/i/os/ ce/rr#a/dos. -> Decassílabo -> 2 6 10

O/ an/ta/go/n#is/mo #e/ra i/ne/vi/t#á/vel. -> Decassílabo -> 5 6 10

Os/ a/ssal/t#an/tes/ #e/ram/ re/pe/l#i/dos. -> Decassílabo -> 4 6 10

Poema 3

O poema é 3 de rima cruzada, métrica decassílaba, duas estrofes e quatro versos, com sílabas “ca” e “to” a serem rimadas, similaridade por esquema rítmico, comparação de versos com o primeiro verso de cada estrofe. Como avaliação o poema recebe 0,81 de fonética, 0,33 de tônicas similares, 0,64 de esquema rítmico e 0,50 de acentuação, gerando um total de 2,29, que corresponde a 57% do valor geral, classificando o poema com boa pontuação.

De repente, uma variante trágica.

Centraliza um horizonte vasto.

Acabaram-se as munições de boca.

menor escala, continua revoltos.

E falseavam a peleja franca.

Anularam-lhes, no fim de algum tempo, o intento.

Quedaram na defensiva franca.

E fora inteiramente imprevisto.

De/ re/p#en/te, #u/ma/ va/ri/#an/te/ tr#á/gi/ca. -> Decassílabo -> 3 4 8 10

Cen/tra/l#i/za/ um/ ho/ri/z#on/te/ v#as/to. -> Decassílabo -> 3 8 10

A/ca/b#a/ram§/se As/ mu/ni/ç#ões/ de/ b#o/ca. -> Decassílabo -> 3 8 10

me/n#or/ es/c#a/la,/ con/ti/n#ua/ re/v#ol/to. -> Decassílabo -> 2 4 8 10

E/ fal/se/#a/vam/ a/ pe/l#e/ja/ fr#an/ca. -> Decassílabo -> 4 8 10

Anularam§/lhes,/ no/ f#im/ de/ al/g#um/ t#em/po, O in/t#en/to. -> Decassílabo -> 4 7 8 10

Qu/e/d#a/ram/ na/ de/fen/s#i/va/ fr#an/ca. -> Decassílabo -> 3 8 10

E/ f#o/ra in/t#e/i/ra/men/te im/pre/v#is/to. -> Decassílabo -> 2 4 10

Anexo

Poema 4

O poema é 4 de rima cruzada, métrica decassílaba, duas estrofes e quatro versos, com sílabas “sa”, “ro”, “to” e “ga” a serem rimadas, similaridade por esquema rítmico, comparação de versos com o primeiro verso de cada estrofe. Como avaliação o poema recebe 1,31 de fonética, 0,33 de tônicas similares, 0,42 de esquema rítmico e 0,50 de acentuação, gerando um total de 2,57, que corresponde a 33% do total geral, classificando o poema com baixa pontuação.

A marcha foi difícil e morosa.
O bombardeio reduziu-se a um tiro.
Uma antinomia vergonhosa.
Assim todo sertanejo é vaqueiro.

Como se extingue o deserto.
Rompe-o e afunda na caatinga.
Entra-se, de surpresa, no deserto.
Distende-se pela orla da caatinga.

A/ m#ar/cha/ f#oi/ di/f#i/cil/ E/ mo/r#o/sa. -> Decassílaba -> 2 4 6 10
O/ bom/bar/d#e/io/ re/du/z#iu\$/se a um/ t#i/ro. -> Decassílaba -> 4 8 10
#U/ma/ an/ti/no/m#i/a/ ver/go/nh#o/sa. -> Decassílaba -> 1 6 10
A/ss#im/ t#o/do/ ser/ta/n#e/jo É/ va/qu#ei/ro. -> Decassílaba -> 2 3 7 10

C#o/mo/ se/ ex/t#in/gu/e/ O/ de/s#er/to. -> Decassílaba -> 1 5 10
R#om/pe\$/o/ E/ a/f#un/da/ na/ caa/t#in/ga. -> Decassílaba -> 1 6 10
#En/tra\$/se,/ de/ sur/pr#e/sa,/ no/ de/s#er/to. -> Decassílaba -> 1 6 10
Dis/t#en/de\$/se/ p#e/la #or/la/ da/ caa/t#in/ga. -> Decassílaba -> 2 5 6 10

Poema 5

O poema 5 é de rima interpolada, métrica dodecassílaba, duas estrofes e quatro versos, com sílabas “ras” e “va” a serem rimadas, similaridade por esquema rítmico, comparação de versos com o primeiro verso de cada estrofe. Como avaliação o poema recebe 1,0 de fonética, 0,33 de tônicas similares, 0,40 de esquema rítmico e 0,50 de acentuação, gerando um total de 2,24, que corresponde a 56% do total geral, classificando o poema com boa pontuação.

O que se segue são vagas conjeturas.
Absorvia-o a psicose coletiva.
Servidão inconsciente; vida primitiva.
Ermavam-se as oficinas e as culturas.

Iludeos com essas iguarias bárbaras.
Começou a erigir-se a igreja nova.
O certo é que os dominava e disciplinava.
Um único às vezes escapava, às carreiras.

O/ que/ se/ s#e/gu/e/ São/ v#a/gas/ con/je/t#u/ras. -> Dodecassílaba -> 4 8 12
Ab/sor/v#i/a\$/o/ a/ psi/c#o/se/ co/le/t#i/va. -> Dodecassílaba -> 3 8 12
Ser/vi/d#ão in/cons/ci/#en/te;/ v#i/da/ pri/mi/t#i/va. -> Dodecassílaba -> 3 6 8 12
Er/m#a/vam\$/se/ As/ o/fi/c#i/nas/ E As/ cul/t#u/ras. -> Dodecassílaba -> 2 8 12

I/lu/d#e/os/ com/ #e/ssas/ i/gua/r#i/as/ b#ár/ba/ras. -> Dodecassílaba -> 3 6 10 12
Co/me/ç#o/u/ a/ e/ri/g#ir\$/se a i/gr#e/ja/ n#o/va. -> Dodecassílaba -> 3 8 10 12
O/ c#er/to É/ que Os/ do/mi/n#a/va E/ dis/ci/pli/n#a/va. -> Dodecassílaba -> 2 7 12
Um/ #ú/ni/co #às/ v#e/zes/ es/ca/p#a/va, #às/ ca/rr#e/i/ras. -> Dodecassílaba -> 2 4 5 9 10 12

Recordações do Escrivão Isaías Caminha

Poema 6

O poema 6 é de rima emparelhada, métrica hendecassílaba, duas estrofes e quatro versos, com sílabas “da” e “me” a serem rimadas, similaridade por esquema rítmico, comparação entre versos de cada estrofe. Como avaliação o poema recebe 0,45 de fonética, 0,25 de tônicas similares, 0,66 de esquema rítmico e 0,50 de acentuação, gerando um total de 1,86, que corresponde a 47% do total geral, classificando o poema com baixa pontuação.

De dia para dia, o jornal crescia em venda.
A redação estava embaçada.
Do interior de um café, o Laje chamou-me.
Era assim sempre que ele falava ao encontrar-me.

Chegou e eu esperei ainda.
Todos nos viramos para a porta de entrada.
Todo o jornal convergia para o crime.
Durante a viagem a angústia avolumou-se-me.

De/ d#ia/ para/ d#i/a, O/ jor/n#al/ cres/c#ia/ Em/ v#en/da. -> Hendecassílaba -> 2 4 7 9 11
A/ re/da/ç#ã/o/ es/t#a/va em/bas/ba/c#a/da. -> Hendecassílaba -> 4 7 11
Do in/te/ri/#or/ de um/ ca/f#é,/ O/ L#a/je/ c#ha/mo/u/me. -> Hendecassílaba -> 4 7 9 11
#E/ra/ a/ss#im/ s#em/pre/ que #e/le/ fa/l#a/va ao e#n/con/trar/me. -> Hendecassílaba -> 1 4 5 7 10 11

Che/g#o/u/ E/ #eu/ es/pe/r#e/i/ a/#in/da. -> Hendecassílaba -> 2 5 8 11
T#o/dos/ nos/ vi/r#a/mos/ para a/ p#or/ta/ de en/tr#a/da. -> Hendecassílaba -> 1 5 8 11
T#o/do/ O/ jor/n#al/ con/ver/g#ia/ para/ O/ cr#i/me. -> Hendecassílaba -> 1 5 8 11
Du/r#an/te a/ vi/#a/gem/ a an/g#ús/ti/a a/#vo/lu/mo/u/se/me. -> Hendecassílaba -> 2 5 8 11

Poema 7

O poema 7 é de rima emparelhada, métrica eneassílaba, duas estrofes e quatro versos, com sílabas “to” e “do” a serem rimadas, similaridade por esquema rítmico, comparação entre versos de cada estrofe. Como avaliação o poema recebe 0,63 de fonética, 0,0 de tônicas similares, 0,66 de esquema rítmico e 0,50 de acentuação, gerando um total de 1,78, que corresponde a 45%, do total geral, classificando o poema com baixa pontuação.

Não serve pra nada, fique certo.
Não consegui realizar tudo isto.
Bem, retrucou o doutor Ricardo.
Trabalharia em quê? em tudo.
A Alfândega ficava perto.
Esperei junto a um café aberto.
Passei o resto do dia vagueando.
Olhou-me com olhar de entendido.

N#ão/ s#er/ve/ pr#a/ n#a/da,/ f#i/que/ c#er/to. -> Eneassílaba -> 1 2 4 5 7 9
N#ão/ con/se/gu#i/ rea/li/z#ar/ t#u/do #is/to. -> Eneassílaba -> 1 4 7 8 9
B#em,/ re/tru/c#ou O/ do/u/t#or/ Ri/c#ar/do. -> Eneassílaba -> 1 4 7 9
Tra/ba/lha/r#i/a / Em/ qu#ê?/ Em/ t#u/do. -> Eneassílaba -> 4 7 9
A/ Al/f#ân/de/ga/ fi/c#a/va/ p#er/to. -> Eneassílaba -> 3 7 9
Es/pe/r#e/i/ j#un/to a um/ ca/f#é/ a/b#er/to. -> Eneassílaba -> 3 4 7 9
Pa/ss#ei O/ r#es/to/ do/ d#ia/ va/gue/#an/do. -> Eneassílaba -> 2 3 6 9
O/#lhou/me/ com/ o/lh#ar/ de en/ten/d#i/do. -> Eneassílaba -> 2 6 9

Contrastes e Confrontos

Poema 8

O poema 8 é de rima cruzada, métrica ocatassílabo, duas estrofes e quatro versos, com sílabas “ta” e “te” a serem rimadas, similaridade por esquema rítmico, comparação entre versos de cada estrofe. Como avaliação o poema recebe 0,80 de fonética, 0,0 de tônicas similares, 0,78 de esquema rítmico e 0,50 de acentuação, gerando um total de 2,08, que corresponde a 52% do total geral, classificando o poema com boa pontuação.

A ilusão é completa.
É um contraste comovente.
A ilusão é completa.
É a região ao sul do Tibete.

Uma obsessão de artista.
É uma sugestão empolgante.
É uma estátua morta.
De longe em longe um caminhante.

A/ i/lu/s#ã/o/ É/ com/pl#e/ta. -> Octassílabo -> 4 8
É/ um/ con/tr#as/te/ co/mo/v#en/te. -> Octassílabo -> 4 8
A/ i/lu/s#ã/o/ É/ com/pl#e/ta. -> Octassílabo -> 4 8
É a/ re/gi/#ão ao/ s#ul/ do/ Ti/b#e/te. -> Octassílabo -> 4 5 8

#U/ma/ ob/se/ss#ã/o/ de ar/t#is/ta. -> Octassílabo -> 1 5 8
É #u/ma/ su/ges/t#ã/o em/pol/g#an/te. -> Octassílabo -> 1 5 8
É/ #u/ma/ es/t#ã/tu/a/ m#or/ta. -> Octassílabo -> 2 5 8
De/ l#on/ge Em/ l#on/ge um/ ca/mi/nh#an/te. -> Octassílabo -> 2 4 8

Poema 9

O poema 9 é de rima cruzada, todas as métricas, duas estrofes e quatro versos, com sílabas “ta” e “so” a serem rimadas, similaridade por esquema rítmico, comparação com primeiro verso de cada estrofe. Como avaliação o poema recebe 1,0 de fonética, 1,0 de tônicas similares, 0,47 de esquema rítmico e 0,50 de acentuação, gerando um total de 2,97, que corresponde a 74% do total geral, classificando o poema com boa pontuação.

A ilusão é completa.
Aplaudem-nos? É duvidoso.
A ilusão é completa.
É um instantâneo prodigioso.

Uma obsessão de artista.
Preferiu o último caso.
É uma estátua morta.
Sustive-a, por isto, um momento, indeciso.

A/ i/lu/s#ã/o/ É/ com/pl#e/ta. -> Octassílabo -> 4 8
A/#plau/dem/nos?/ É/ du/vi/d#o/so. -> Octassílabo -> 2 8
A/ i/lu/s#ã/o/ É/ com/pl#e/ta. -> Octassílabo -> 4 8
É um/ ins/tan/t#â/neo/ pro/di/gi/#o/so. -> Eneassílabo -> 4 9
#U/ma ob/se/ss#ã/o/ de ar/t#is/ta. -> Heptassílabo -> 1 4 7
Pre/fe/r#i/u O #úl/ti/mo/ c#a/so. -> Heptassílabo -> 3 4 7
É/ #u/ma es/t#ã/tu/a/ m#or/ta. -> Heptassílabo -> 2 4 7
#Sus/ti/ve/a,/ por/ #is/to,/ um/ mo/m#en/to, in/de/c#i/so. -> Barbáro -> 1 6 10 13

Anexo

Poema 10

O poema 10 é de rima cruzada, todas as métricas, duas estrofes e quatro versos, com sílabas “ta” e “so” a serem rimadas, similaridade por esquema rítmico, comparação com primeiro verso de cada estrofe. Como avaliação o poema recebe 1,06 de fonética, 1,0 de tônicas similares, 0,54 de esquema rítmico e 0,50 de acentuação, gerando um total de 3,10, que corresponde a 77% do total geral, classificando o poema com boa pontuação.

A ilusão é completa.

É um instantâneo prodigioso.

A ilusão é completa.

Aplaudem-nos? É duvidoso.

É uma estátua morta.

Preferiu o último caso.

Lança o olhar em volta.

Sustive-a, por isto, um momento, indeciso.

A/ i/lu/s#ã/o/ É/ com/pl#e/ta. -> Octossílabo -> 4 8

É um/ ins/tan/t#â/ne/o/ pro/di/gi/#o/so. -> Decassílabo -> 4 10

A/ i/lu/s#ã/o/ É/ com/pl#e/ta. -> Octossílabo -> 4 8

A/#plau/dem/nos?/ É/ du/vi/d#o/so. -> Octossílabo -> 2 8

É #u/ma es/t#á/tu/a/ m#or/ta. -> Hexassílabo -> 1 3 6

Pre/fe/r#iu O #úl/ti/mo/ c#a/so. -> Hexassílabo -> 3 6

L#an/ça/ O o/lh#ar/ Em/ v#ol/ta. -> Hexassílabo -> 1 4 6

#Sus/ti/ve/a,/ por/ #is/to, um/ mo/m#en/to, in/de/c#i/so. -> Dodecassílabo -> 1 6 9 12

Os Lusíadas

Poema 11

O poema 11 é de rima cruzada, métrica decassílabo, duas estrofes e quatro versos, com sílabas “nos” e “d#eu” a serem rimadas, similaridade por esquema rítmico, comparação com primeiro verso de cada estrofe. Como avaliação o poema recebe 1,89 de fonética, 0,33 de tônicas similares, 0,81 de esquema rítmico e 0,50 de acentuação, gerando um total de 3,53, que corresponde a 88% do total geral, classificando o poema com boa pontuação.

A senhores em tudo soberanos.

Manuel, que a Joane sucedeu.

Portugal, mas os coros soberanos.

Sobre a terra Africana descendeu.

Mas o Mouro, instruído nos enganos.

Mais ladrões, castigando, à morte deu.

Que por ela se esqueçam os humanos.

Ao grande Saladino se rendeu.

A/ se/nh#o/res/ Em/ t#u/do/ so/be/r#a/nos. -> Decassílabo -> 3 6 10

Ma/nu/#el,/ que a/ Jo/#a/ne/ su/ce/d#eu. -> Decassílabo -> 3 6 10

Por/tu/g#al,/ Mas/ Os/ c#o/ros/ so/be/r#a/nos. -> Decassílabo -> 3 6 10

S#o/bre a/ t#e/rra A/fri/c#a/na/ des/cen/d#eu. -> Decassílabo -> 1 3 6 10

Mas/ O/ M#ou/ro, ins/tru/#i/do/ nos/ en/g#a/nos. -> Decassílabo -> 3 6 10

M#ais/ la/dr#ões,/ cas/ti/g#an/do, à/ m#or/te/ d#eu. -> Decassílabo -> 1 3 6 8 10

Que/ por/ #e/la/ se es/qu#e/çam/ Os/ hu/m#a/nos. -> Decassílabo -> 3 6 10

Ao/ gr#an/de/ Sa/la/d#i/no/ se/ ren/d#eu. -> Decassílabo -> 2 6 10

Anexo

Poema 12

O poema 12 é de rima cruzada, métrica bárbara, duas estrofes e quatro versos, com sílabas “ma” e “so” a serem rimadas, similaridade por esquema rítmico, comparação com primeiro verso de cada estrofe. Como avaliação o poema recebe 1,0 de fonética, 0,0 de tônicas similares, 0,62 de esquema rítmico e 0,50 de acentuação, gerando um total de 2,12, que corresponde a 53% do total geral, classificando o poema com boa pontuação.

À Moçambique esta Ilha, que se chama.
Pouco val coração, astúcia e siso.
Nas cousas de que tem notícia e fama.
Quando os Deuses no Olimpo luminoso.

Com as argêntas caudas branca escuma.
Que assi vai alternando o tempo iroso.
E Judeia, que um Deus adora e ama.
Perseguem-no com as lanças, e ele, iroso.

À/ Mo/çam/b#i/qu/e/ #es/ta/ #I/lha,/ que/ se/ ch#a/ma. -> Barbáro -> 4 7 9 13
P#o/u/co/ v#al/ co/ra/ç#ão,/ as/t#ú/ci/a/ E/ s#i/so. -> Barbáro -> 1 4 7 9 13
Nas/ c#o/u/sas/ de/ que/ t#em/ no/t#í/ci/a/ E/ f#a/ma. -> Barbáro -> 2 7 9 13
Qu#an/do/ Os/ D#e/u/ses/ no/ O/l#im/po/ lu/mi/n#o/so. -> Barbáro -> 1 4 9 13

Com/ As/ ar/g#ên/te/as/ c#a/u/das/ br#an/ca/ es/c#u/ma. -> Barbáro -> 4 7 10 13
Que/ a/ss#i/ v#ai/ al/ter/n#an/do/ O/ t#em/po/ i/r#o/so. -> Barbáro -> 3 4 7 10 13
E/ Ju/d#e/i/a,/ que/ um/ D#eus/ a/d#o/ra/ E/ #a/ma. -> Barbáro -> 3 8 10 13
P#er/se/guem/no/ com/ As/ l#an/ças,/ E/ #e/le,/ i/r#o/so. -> Barbáro -> 1 7 10 13

Poema 13

O poema 13 é de rima emparelhada, métrica hendecassílaba, duas estrofes e quatro versos, com sílabas “t#iu” e “o” a serem rimadas, similaridade por esquema rítmico, comparação com primeiro verso de cada estrofe. Como avaliação o poema recebe 1,17 de fonética, 0,75 de tônicas similares, 0,70 de esquema rítmico e 0,50 de acentuação, gerando um total de 3,12, que corresponde a 78% do total geral, classificando o poema com boa pontuação.

Logo cada um dos Deuses se partiu.
Logo todo o restante se partiu.
Tétis todo o cerúleo senhorio.
Que onde reina a malícia, está o receio.

Muitos, que Deus o quis e permitiu.
A cabeça inclinando, consentiu.
Lava, e o Danúbio, Amásis e Álbis rio.
Que agora tem do mundo o senhorio.

L#o/go/ c#a/da um/ dos/ D#e/u/ses/ se/ par/t#iu. -> Hendecassílabo -> 1 3 6 11
L#o/go/ t#o/do/ O/ res/t#an/te/ se/ par/t#iu. -> Hendecassílabo -> 1 3 7 11
T#é/tis/ t#o/do O/ ce/r#ú/le/o/ se/nho/r#i/o. -> Hendecassílabo -> 1 3 6 11
Que #on/de/ r#ei/na a/ ma/l#í/cia, es/t#á/ O/ re/c#ei/o. -> Hendecassílabo -> 1 3 6 8 11

M#u/i/tos,/ que/ D#eus/ O/ qu#is/ E/ per/mi/t#iu. -> Hendecassílabo -> 1 5 7 11
A/ ca/b#e/ça/ in/cli/n#an/do,/ con/sen/t#iu. -> Hendecassílabo -> 3 7 11
L#a/va,/ E O/ Da/n#ú/bio, A/m#á/sis/ E #ál/bis/ r#i/o. -> Hendecassílabo -> 1 5 7 9 11
Que/ a/g#o/ra/ t#em/ do/ m#un/do O/ se/nho/r#i/o. -> Hendecassílabo -> 3 5 7 11