

UNIVERSIDADE FEDERAL DO PARANÁ

ANTONIO WILLIAN SOUSA

AVALIAÇÃO E VALORAÇÃO AUTOMÁTICA DA COESÃO DE TEXTOS DISSERTATIVOS

CURITIBA PR

2021

ANTONIO WILLIAN SOUSA

AVALIAÇÃO E VALORAÇÃO AUTOMÁTICA DA COESÃO DE TEXTOS DISSERTATIVOS

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Informática no Programa de Pós-Graduação em Informática, Setor de Ciências Exatas, da Universidade Federal do Paraná.

Área de concentração: *Ciência da Computação*.

Orientador: Marcos Didonet del Fabro.

CURITIBA PR

2021

Catálogo na Fonte: Sistema de Bibliotecas, UFPR
Biblioteca de Ciência e Tecnologia

S725a

Sousa, Antonio Willian

Avaliação e valoração automática da coesão de textos dissertativos
[recurso eletrônico] / Antonio Willian Sousa. – Curitiba, 2021.

Dissertação - Universidade Federal do Paraná, Setor de Ciências Exatas,
Programa de Pós-Graduação em Informática, 2021.

Orientador: Marcos Didonet del Fabro

1. Processamento da linguagem natural (Computação). 2. Coesão
(Linguística). 3. Processamento de textos (Computação). 4. Análise do
discurso. I. Universidade Federal do Paraná. II. Didonet del Fabro, Marcos .
III. Título.

CDD: 006.35

Bibliotecário: Elias Barbosa da Silva CRB-9/1894



MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIÊNCIAS EXATAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO INFORMÁTICA -
40001016034P5

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em INFORMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **ANTONIO WILLIAN SOUSA** intitulada: **Avaliação e Valoração Automática da Coesão de Textos Dissertativos**, sob orientação do Prof. Dr. MARCOS DIDONET DEL FABRO, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 13 de Agosto de 2021.

Assinatura Eletrônica

13/08/2021 14:10:25.0

MARCOS DIDONET DEL FABRO

Presidente da Banca Examinadora

Assinatura Eletrônica

13/08/2021 14:43:12.0

FABIANO SILVA

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

13/08/2021 15:36:20.0

HEGLER CORREA TISSOT

Avaliador Externo (UNIVERSIDADE DREXEL)

Aos meus pais Maria da Graças e Jonas Marques, pois sem eles nada disso seria possível.

Aos meus irmãos Sandra e Aucélio, pelo seu amor, por sua ajuda e por serem exemplos para mim.

À minha esposa Monique por todo o companheirismo, amor e suporte em todos os momentos.

Ao meu filho Álix que desde o seu nascimento, a cada dia me incentiva e me ensina a ser uma pessoa melhor.

AGRADECIMENTOS

Agradeço ao meu orientador Marcos Didonet del Fabro pela sua parceria e por servir como um guia nos momentos mais difíceis, compartilhando o seu conhecimento e experiência.

Agradeço a todos que direta ou indiretamente fizeram parte desta história e contribuíram para a realização deste trabalho.

RESUMO

Em um texto escrito, a representação das ideias e das informações que se deseja transmitir devem estar organizadas de tal maneira que permita ao leitor receber, perceber e decodificar a informação de maneira lógica, sequencial e, em última instância, apreensível. Esta organização é responsável pelo estabelecimento, por meio do texto, de um canal de comunicação entre quem transmite e quem recebe a informação. A construção deste canal estabelecido entre escritor e leitor necessita de um suporte que é construído à medida que a leitura ocorre. Essa construção se sustenta nos mecanismos e métodos de coesão textual, que são responsáveis por aquilo que se pode qualificar como um bom texto e que dão unidade a ele, realizando a conexão de ideias de uma maneira lógica. Dado um texto escrito, avaliar a qualidade textual, legibilidade, coerência e o seu poder de transmitir a informação correta, certamente passa pela avaliação do bom uso dos mecanismos de coesão. Estes mecanismos não se constituem apenas de características gramaticais e sintáticas, pois além de atributos deste tipo, também permeiam e agregam informações e recursos semânticos e pragmáticos. Exigem assim, que métodos de avaliação de coesão textual - inclusive os automáticos - apresentem robustez suficiente para capturar informações e conexões dos diferentes aspectos da língua escrita. Os trabalhos de valoração de textos encontrados na literatura relacionada, em geral, não abordam dimensões específicas. O objetivo destes trabalhos é fornecer uma pontuação global para o texto, sem fornecer um valor para cada competência específica. Todavia, algumas propostas de análise a partir de dimensões como a coesão, coerência e adequação ao tema podem ser encontrados, mas estas propostas se utilizam do mesmo conjunto de atributos ou do mesmo modelo para avaliar aspectos diferentes do texto. Além disso, os atributos utilizados são considerados superficiais, pois não estão baseados em teorias linguísticas e também não permitem uma análise mais profunda da estruturação textual. Dessa forma, com uso das referidas propostas de avaliação não é possível prover um retorno acerca da qualidade, da relação e da interconexão de elementos estruturantes que são essenciais para um texto. Neste trabalho propõe-se um método de valoração da coesão textual baseado na Teoria da Centralidade e na Teoria do Foco que pode ser utilizado na avaliação de redações, na melhoria da escrita de alunos e no suporte de processos educacionais. Para isso, foram feitas integrações entre diferentes ferramentas de correção gramatical e ortográfica, menção de entidades, resolução de correferências e o treinamento de um modelo de vinculação de entidades, específico para o português, baseado em dados da Wikipedia. Para validar o método escolhido, construiu-se uma base de redações contendo as valorações das competências esperadas dos seus autores e desenvolveu-se um protótipo que implementa o método proposto. Aplicou-se o protótipo desenvolvido sobre um conjunto de redações e os valores obtidos relacionados à competência da coesão textual foram comparados com as notas de avaliadores humanos e com os resultados de trabalhos correlatos. Os resultados obtidos mostram o potencial desta abordagem, pois obteve-se resultados tão bons quanto os estudos considerados o estado-da-arte, porém com um ganho de tempo considerável e com um esforço muito menor de obtenção de informações e atributos do textos, sendo possível seu uso tanto na melhoria de sistemas de valoração de redações, quanto no desenvolvimento de ferramentas de apoio a processos educacionais formativos relacionados ao desenvolvimento da escrita.

Palavras-chave: processamento de linguagem natural. coesão textual. avaliação automática de textos. teoria do discurso. teoria da centragem. coesão local. coesão global.

ABSTRACT

In a written text, the representation of ideas and information to be transmitted must be organized in such a way that allows the reader to receive, perceive and decode the information in a logical, sequential and, most importantly, apprehensible manner. Such organization is responsible for establishing, through the text, a communication channel between the one who transmits and the one who receives the information. The construction of this channel, established between writer and reader, needs a support that is built as the reading happens. This construction is supported by the mechanisms and methods of textual cohesion, which are responsible for what can be qualified as a good text and which give unity to it, making the connection of ideas in a logical way. Given a written text, evaluating its textual quality, its readability, its coherence and its power to communicate the correct information, certainly involves evaluating the good use of the mechanisms of cohesion. These mechanisms do not consist only of grammatical and syntactic features, because in addition to attributes of this kind, they also permeate and aggregate semantic and pragmatic information and resources. Thus, methods for evaluating textual cohesion - including automatic ones - must be robust enough to capture information and connections of the different aspects of the written language. The text evaluation papers found in the related literature generally do not address specific dimensions of texts. The goal of these works is to provide an overall score for the text, without providing a value for each specific competence. However, some proposals for analysis based on dimensions such as cohesion, coherence and adequacy to the topic can be found, but these proposals use the same set of attributes or the same model to evaluate different aspects of the text. Moreover, the attributes used are considered superficial, as they are not based on linguistic theories and also do not allow a deeper analysis of textual structuring. Thus, using these evaluation proposals it is not possible to provide feedback about the quality, the relationship and the interconnection of structuring elements that are essential to a text. In this paper, we propose a method for evaluating textual cohesion, based on the Centrality Theory and the Focus Theory, that can be used to evaluate essays, to improve students' writing, and to support educational processes. To this end, integrations between different tools for grammar and spelling correction, entity mentioning, coreference resolution, and the training of a Portuguese-specific model of entity linking based on Wikipedia data were performed. To validate the chosen method, an essay database was built containing the scores of the competencies required for dissertative essays, and a prototype was developed that implements the proposed method. The prototype was applied to these essays and the values obtained were compared to the scores of human evaluators and similar works regarding textual cohesion. The results obtained show the potential of this approach, because it obtained results as good as the studies considered state-of-the-art, but with a considerable gain in time and with much less effort to obtain information and text attributes. It is therefore possible to use it to improve writing evaluation systems, as well as to develop tools to support educational and formative processes related to the development of writing.

Keywords: natural language processing. textual cohesion. text automatic evaluation. discourse theory. centering theory. local cohesion. global cohesion.

LISTA DE FIGURAS

2.1	Estruturação do processo de análise de coesão.	23
2.2	Processo wikification.. . . .	29
3.1	Arquitetura do Intelligent Essay Assessor.	37
3.2	Tela de resultados do Intelligent Essay Assessor.. . . .	37
4.1	Avaliação e valoração de coesão.	45
5.1	Distribuição das notas da Competência 4.	62
5.2	Comparação dos valores atribuídos pelo avaliador humano (esquerda) com os valores de coesão calculada (direita).. . . .	64
5.3	Distribuição das diferenças absolutas das notas.	65
5.4	Comparação dos valores atribuídos pelo avaliador humano (esquerda) com os valores de coesão calculada pelo método proposto no Experimento 2 (direita). . .	67
5.5	Comparação dos valores de coesão calculados no Experimento 1 (esquerda) e Experimento 2 (direita) pelo método proposto.	67
5.6	Valores da coesão local dos pares de sentenças calculados pelo método.	73
5.7	Valores da coesão global dos pares de parágrafos calculados pelo método.	73
5.8	Exemplo de texto contendo apenas transições de elaboração.	74
5.9	Exemplo de texto com variações nas relações entre sentenças.. . . .	74
B.1	Valores da coesão local dos pares de sentenças.	88
B.2	Valores da coesão local dos pares de parágrafos.. . . .	88

LISTA DE TABELAS

2.1	Relação entre FE e FI para estabelecimento da coesão e valoração.	25
2.2	Exemplo de avaliação da coesão para duas sentenças.	25
2.3	Corpus de Redações para o Português Brasileiro.	31
3.1	Grupos de índices utilizados no TAACO.	35
3.2	Sistemas de valoração automática de redações.. . . .	38
3.3	Dimensões de análise da qualidade textual.. . . .	39
3.4	Abordagens de redes neurais e modelagem de documentos.	40
3.5	Grupos de atributos utilizados.	41
4.1	Anotação de texto usando a biblioteca Spacy.	47
4.2	Anotação de texto usando o parser PALAVRAS.. . . .	48
4.3	Anotação de uma entidade usando DBpedia Spotlight.. . . .	48
4.4	Exemplos de erros do CoGrOO.	50
4.5	Exemplo de Entradas da Base de Conhecimento.. . . .	51
4.6	Entidade com duas representações similares.. . . .	51
4.7	Lista de regras do CORP.	52
4.8	Texto separado em sentenças..	53
4.9	Tabela de Cadeias de Correferências.	54
4.10	Anotação de Informações das Sentenças	56
4.11	Valor da Relação dos Pares de Sentenças	57
4.12	Avaliação dos Pares de Parágrafos	58
4.13	Resultados da Avaliação dos Pares de Parágrafos	59
5.1	Dados do conjunto de redações.	62
5.2	Valores de média, desvio-padrão e distribuição em quartis dos valores das notas dada pelo avaliador humano e das notas calculadas pelo método nos Experimentos 1 e 2.. . . .	66
5.3	Valores de QWK, quanto maior, melhor. Valores de RMSE, quanto menor, melhor. (1) Resultados utilizando o método de Amorim e Veloso. (2) Resultados utilizando o método de Fonseca et al..	68
5.4	Resultado da avaliação de diferentes modelos de regressão, ordenados do menor para o maior.	70
5.5	Valores de QWK, quanto maior, melhor. Valores de RMSE, quanto menor, melhor. (1) Resultados utilizando o método de Amorim Veloso. (2) Resultados utilizando o método de Fonseca et al.	70

5.6	Resultados do Experimento 3 sem correferências e entidades..	70
5.7	Relação entre FE e FI para estabelecimento da coesão e valoração.	71
5.8	Texto separado em sentenças..	72
A.1	Tabela de Erros Gramaticais.	81
A.2	Erros de paronímia, regência e repetição..	83
B.1	Tabela de Cadeias de Correferências.	87
C.1	Lista de regras do CORP.	101

LISTA DE ACRÔNIMOS

PLN	Processamento de Linguagem Natural
RST	Rhetorical Structure Theory
AAR	Análise Automática de Redações
AAD	Análise Automática Discursiva
VAR	Valoração Automática de Redações
TOEFL	Test of English as a Foreign Language
GRE	Graduate Record Examinations
CET	College English Test
ENEM	Exame Nacional do Ensino Médio
MEC	Ministério da Educação e Cultura
IC	Índice de Coesão
IF	Índice de Forma
QWK	Quadratic Weighted Kappa
RMSE	Root Mean Square Error
TAACO	Tool for the Automatic Analysis of Cohesion
PEG	Project Essay Grader
IEA	Intelligent Essay Assessor
LSA	Latent Semantic Analysis
MAC	Módulo de Análise de Coerência
SciPo	Scientific Portuguese
REN	Resolução de Entidades Nomeadas
UD	Universal Dependencies
FI	Foco Implícito
FE	Foco Explícito
RDF	Resource Description Framework
URI	Uniform Resource Identifier

SUMÁRIO

1	INTRODUÇÃO	14
1.1	MOTIVAÇÃO.	14
1.2	DESAFIOS	16
1.3	HIPÓTESE	17
1.4	OBJETIVOS	17
1.5	CONTRIBUIÇÃO	17
1.6	ORGANIZAÇÃO DO DOCUMENTO	18
2	FUNDAMENTAÇÃO TEÓRICA.	19
2.1	PROCESSAMENTO DE LINGUAGEM NATURAL	19
2.2	COESÃO E COERÊNCIA TEXTUAL	19
2.2.1	Teoria do Discurso	20
2.2.2	Coesão Textual	21
2.2.3	Coerência Textual	26
2.3	RESOLUÇÃO DE CORREFERÊNCIAS E VINCULAÇÃO DE ENTIDADES	27
2.3.1	Resolução de Correferências	27
2.3.2	Detecção e Vinculação de Entidades	29
2.4	AVALIAÇÃO AUTOMÁTICA DE REDAÇÕES.	30
2.5	AVALIAÇÕES E MÉTRICAS	31
2.6	CONCLUSÃO	33
3	REVISÃO BIBLIOGRÁFICA	34
3.1	AVALIAÇÃO DE COESÃO E COERÊNCIA	34
3.2	AVALIAÇÃO AUTOMÁTICA DE REDAÇÕES.	36
3.3	VALORAÇÃO DE REDAÇÕES EM PORTUGUÊS.	40
3.4	DISCUSSÃO	43
3.5	CONCLUSÃO	43
4	VALORAÇÃO E AVALIAÇÃO AUTOMÁTICA DA COESÃO	45
4.1	MÉTODO DE VALORAÇÃO DA COESÃO.	45
4.2	ETIQUETAGEM SEMÂNTICA E SINTÁTICA.	47
4.3	DICIONÁRIOS E BASES DE SINÔNIMOS	49
4.4	CORRETOR GRAMATICAL	49
4.5	VINCULAÇÃO DE ENTIDADES	50
4.6	RESOLUÇÃO DE CORREFERÊNCIA	51
4.6.1	Cálculo da Coesão	54
4.6.2	Comparação com Outros Trabalhos	59

4.7	CONCLUSÃO	60
5	EXPERIMENTOS E RESULTADOS	61
5.1	BASE DE TEXTOS.	61
5.2	EXPERIMENTOS	63
5.2.1	Experimento 1: Valoração com Entidades e Referências.	63
5.2.2	Experimento 2: Valoração com Uso de Correferência	65
5.2.3	Experimento 3: Valoração com Índice de Forma e de Imprecisão	68
5.3	COESÃO PARA APOIO DE AVALIAÇÕES FORMATIVAS.	71
5.4	CONCLUSÃO	75
6	CONCLUSÃO	76
6.1	CONCLUSÕES GERAIS.	76
6.2	TRABALHOS FUTUROS	77
	REFERÊNCIAS	78
	APÊNDICE A – ERROS GRAMATICAIS CONSIDERADOS	81
	APÊNDICE B – TEXTOS COM ANOTAÇÕES E DADOS	84
B.1	SENTENÇAS, LISTAS DE FOCO E ENTIDADES	84
B.2	CADEIAS DE CORREFERÊNCIA	87
B.3	COESÃO LOCAL E GLOBAL.	88
B.4	ERROS DETECTADOS PELO COGROO.	88
B.5	ENTIDADES, ALIAS E CANDIDATOS	89
B.6	ANOTAÇÕES PARSER PALAVRAS	90
	APÊNDICE C – TIPOS DE CORREFERÊNCIAS DO CORP	101

1 INTRODUÇÃO

Neste capítulo apresenta-se a motivação para o desenvolvimento deste estudo, os desafios enfrentados, a hipótese relacionada, os objetivos e a contribuição esperada. Ao final, apresenta-se a organização dos demais capítulos do documento.

1.1 MOTIVAÇÃO

A coesão textual é um elemento estruturante presente em todo tipo de texto, desde aqueles utilizados para repasse simples e factual de informações, sem contudo estabelecer um encadeamento de ideias ou linha de raciocínio determinada, até aqueles direcionados para a transmissão de ideias complexas. Segundo Koch e Travaglia (1989), a coesão está intimamente ligada à boa formação textual em termos da interlocução comunicativa. Indiferente do tipo de texto, a correta transmissão de informação por meio daquilo que pode ser considerado como um bom texto é feita ao se estabelecer um canal de comunicação dinâmico que é sustentado pela coesão textual.

Por ser a coesão um elemento essencial do texto bem escrito, ainda que não seja o único, a sua adequada avaliação se mostra essencial em subáreas do Processamento de Linguagem Natural (PLN) como a Análise Automática Discursiva (AAD) e a Avaliação Automática de Redações (AAR). Sendo que esta última, objetiva a criação de sistemas de correção e avaliação automática de textos. Estes sistemas apresentam vantagens em relação ao custo e tempo de execução quando comparadas com corretores humanos em tarefas similares, pois corretores humanos podem discordar em aspectos subjetivos do texto, bem como estão sujeitos ao desgaste físico ao avaliar uma grande quantidade de textos similares de maneira reiterada. Desse modo, os sistemas de correção automática, podem viabilizar uma valoração mais assertiva.

Na literatura relacionada é possível encontrar estudos que buscam demonstrar a viabilidade de sistemas de avaliação automática, baseando-se no fato que os valores dos corretores humanos apresentam divergências e os sistemas automáticos poderiam ser utilizados para minimizá-las. Como um dos exemplos deste tipo trabalho, pode-se citar (Shermis, 2014) que defende que ao considerarmos conteúdo, argumentação e organização é possível um trabalho conjunto entre os avaliadores humanos e os sistemas automáticos. Em contrapartida, trabalhos como (Perelman, 2014) apresentam duras críticas aos sistemas automáticos de valoração. Em especial, são feitas críticas ao fato de serem usadas as comparações entre as notas de humanos e sistemas automáticos para justificar o uso destes últimos, quando os dois - humano e o computador - adquirem os dados de forma distinta. Assim, alguns autores sustentam que os métodos automáticos se baseiam apenas em elementos superficiais ou mecânicos.

A despeito das críticas e da qualidade já alcançada por métodos de valoração automática e sistemas de VAR, há diversos testes aplicados em larga escala, que exigem a escrita de textos e claro, necessitam de métodos de avaliação rápidos. Como exemplo, temos o TOEFL¹ que é aplicado para avaliação do conhecimento da língua inglesa e o GRE², que avalia diversas competências. Se considerarmos a língua chinesa, têm-se o CET³, que é um exame aplicado a

¹O Test of English as a Foreign Language (TOEFL) é um teste padronizado que mede a proficiência na língua inglesa de falantes não-nativos.

²O Graduate Record Examinations (GRE) é um teste padronizado que é pré-requisito para entrar em diversas universidades nos Estados Unidos, Canadá e outros países.

³O College English Test é um teste de inglês como língua estrangeira.

todos os estudantes chineses em processo de graduação. No Brasil, há o Exame Nacional de Desempenho de Estudantes (ENADE) aplicado aos alunos de nível superior e o Exame Nacional do Ensino Médio (ENEM), aplicado aqueles que concluem o ensino médio e desejam ingressar em uma universidade.

A avaliação de redações em exames como o ENEM é feita por meio de uma nota geral, composta por 05 notas que abordam competências distintas. A avaliação destas competências, de acordo com o Ministério da Educação e Cultura (MEC) - responsável pela aplicação do exame - mede a capacidade do autor da redação com relação aos seguintes quesitos:

- conhecimento da norma culta da língua portuguesa;
- compreensão da proposta de redação e aplicação de conceitos de diferentes áreas do conhecimento no desenvolvimento do texto, mantendo o escopo de um texto dissertativo-argumentativo;
- seleção, organização, estabelecimento de relações e interpretação de informações, fatos, opiniões, argumentos para estabelecer a defesa de um ponto de vista;
- conhecimento dos mecanismos linguísticos necessários para a construção da argumentação;
- elaboração de uma proposta de solução para o problema abordado.

Cada uma das competências é avaliada e recebe uma pontuação em uma escala de 0 a 200. Estas competências são somadas para que ao final se tenha a nota global que valoriza a qualidade geral do texto. O processo de avaliação de competências e atribuição das notas devem ser acompanhados de um retorno sobre a percepção do avaliador e as razões pelas quais ele atribuiu determinada nota para cada uma das competências. Avaliações deste tipo que dão um retorno dos pontos fortes e fracos de um texto, são de grande utilidade em avaliações de redações de exames, mas também o são em processos que envolvem avaliação educacional formativa (Santos e Kroeff, 2018), tão importantes nos níveis mais básicos da educação e com potencial para contribuir para uma melhoria da qualidade da educação.

Segundo Santos et al. (2018), a avaliação formativa compreende um acompanhamento realizado de forma continuada e individualizada, o que exige do educador uma quantidade maior de análises e avaliações dos produtos gerados pelo aluno durante as etapas do seu aprendizado. Além disso, a avaliação é contínua e não delimitada em etapas como ocorre em avaliações somativas. Para que a avaliação formativa produza os resultados esperados, é necessário que a velocidade dessas avaliações seja rápida, sem perder qualidade e que, ao mesmo tempo, se entregue ao aluno um retorno compreensível sobre como ele pode progredir na tarefa avaliada, provendo suporte para que o mesmo se torne protagonista da mudança.

A implementação efetiva de avaliações formativas na educação básica brasileira pode enfrentar fortes barreiras relacionadas à velocidade das avaliações e do retorno provido. Segundo o MEC, que é responsável por determinar as diretrizes da educação, em 2020 o Brasil contava com 47,9 milhões de alunos na educação básica e desse total, 38,7 milhões de alunos estão em escolas públicas⁴. De forma contrastante, ainda em 2020, o número de professores da educação básica era de 2,2 milhões. Uma proporção entre alunos e professores, dessa magnitude, dificulta a aplicação de processos formativos, que requerem análises contínuas e frequentes. Considerando esta proporção entre professores e alunos, os métodos automatizados podem se mostrar úteis, em

⁴<http://dadosabertos.mec.gov.br> (Acessado em 01/06/2021)

especial no apoio de processos formativos que exijam análises mais criteriosas, como a avaliação de textos.

De acordo com Shermis e Burstein (2013), os métodos automáticos podem ser de grande valia para avaliações em sala de aula ou no desenvolvimento de habilidades específicas, podendo ser utilizados para valorar os processos de avaliação formativa, ajudar professores e alunos a identificar e direcionar problemas específicos na escrita, por meio da valoração obtida; de forma distinta da avaliação somativa que foca no quanto o estudante aprendeu em um determinado período ou na busca de um objetivo específico. Este tipo de avaliação está naturalmente inclinado a ser feito por um avaliador humano, dado o caráter subjetivo e cumulativo da aprendizagem. Ainda que os próprios Shermis e Burnstein, indiquem que o uso de sistemas de avaliação automática possa ser considerado no futuro como forma de prover *feedbacks* aos estudantes e citando Grimes e Warschauer (2010), defendem que um retorno objetivo pode servir como uma boa forma de incentivar uma melhora em certos aspectos da escrita, como é o caso da coesão.

Os sistemas de avaliação automática se mostram bons candidatos para prover informações sobre a coesão, uma vez que esta é uma propriedade objetiva do texto, construída por meio do uso de palavras, sentenças e outros elementos que conduzem o leitor na interpretação e conexão de ideias, temas e tópicos. Contudo, o uso de avaliadores automatizados em textos do português brasileiro encontrados na literatura, de acordo com o estudo conduzido para embasar este trabalho, mostrou ser comum abordar a coesão como um requisito da coerência, valorando a coesão de forma implícita. Na perspectiva que se pretende apresentar com esta dissertação, ainda que não possa ser considerada incorreta, pode não permitir explorar de forma adequada os aspectos intrínsecos dos quesitos necessários à obtenção de um texto coeso.

Segundo Koch e Travaglia, a coerência difere da coesão, pois está relacionada à representação mental do conteúdo do texto que o leitor cria, apresentando um caráter mais subjetivo e que varia de acordo com nível de conhecimento do leitor acerca do domínio no qual o texto lido se insere. A criação dessa representação mental é feita pelas indicações dadas pelos mecanismos de coesão que mostram ao leitor como montá-la de maneira coerente. Apesar da importância da coesão em conduzir a criação da representação do texto para o leitor, acredita-se que a mesma não tenha sido devidamente explorada pelos métodos de valoração de textos em português.

Neste trabalho será apresentada uma solução para a avaliação e valoração da coesão, com a adoção e adaptação do modelo proposto em (Nobre, 2011), integrado com diversas outras propostas de tarefas de PLN que auxiliam na identificação de elementos importantes na estruturação da coesão. Com este modelo obteve-se boa cobertura das relações responsáveis pelo estabelecimento da coesão, sendo uma abordagem voltada especificamente para a valoração de textos em português. Acredita-se que a avaliação automática de textos, com a devida valoração da estruturação, encadeamento e ordenação dos elementos coesivos, necessita de um método que, diferentemente de imitar a forma de trabalho de um corretor humano ou de utilizar atributos que não tem relação direta com os aspectos textuais analisados, se valha de técnicas baseadas em teorias linguísticas comprovadas. Assim, métodos com estas características, poderão atuar como suporte dos métodos utilizados por corretores humanos em ações educacionais formativas e na facilitação do seu trabalho.

1.2 DESAFIOS

Neste trabalho tem-se os seguintes desafios a serem superados:

- Criação de uma base aberta de textos dissertativos em português brasileiro que permita valorar a coesão textual;

- Integração de sistemas linguístico-computacionais especializados em tarefas de PLN, como correção ortográfica, gramatical, menção, vinculação, e resolução de entidades, dentre outras necessárias para a identificação e avaliação dos elementos, relações e mecanismos responsáveis pela coesão textual;
- Implementação de um sistema de valoração, baseado em um método que permita valorar a qualidade da coesão textual de forma equivalente a um corretor humano;
- Implementação de um sistema que permita avaliar a coesão de um texto oferecendo um retorno sobre a sua estruturação dos mecanismos de coesão.

1.3 HIPÓTESE

Nossa hipótese é que, por meio do uso de recursos e ferramentas existentes, ou que venham a ser criadas no contexto deste trabalho, seja possível valorar a qualidade de textos quanto à sua coesão, obtendo valores semelhantes aqueles obtidos por um avaliador humano e provendo informações de como a valoração foi obtida.

1.4 OBJETIVOS

O objetivo primário será a integração e adaptação de diferentes propostas de resolução de tarefas de PLN visando calcular a qualidade coesiva de um texto em Português brasileiro, sem necessitar da anotação de textos por especialistas e baseando-se em informações que possam ser obtidas a partir do próprio texto e de suas estruturas ortográficas, gramaticais e semânticas.

Como objetivos secundários tem-se a adaptação de um método de cálculo de coesão baseado na Teoria da Centragem, a criação de uma base de textos contendo valorações específicas em relação à qualidade coesiva e, por último, a proposição de um método de inspeção visual da coesão local e global de textos.

1.5 CONTRIBUIÇÃO

Este trabalho traz as seguintes contribuições:

- Análise e desenvolvimento de um protótipo de valoração da coesão textual;
- Visão integrada da estrutura do processo de avaliação da qualidade coesiva de textos dissertativos;
- Disponibilização de uma base aberta de textos dissertativos;
- Proposição de um método de inspeção visual da coesão textual;
- Disponibilização de um conjunto de recursos linguísticos abertos, de uso livre e expansível⁵.

⁵Os dados e códigos utilizados estão disponíveis em <https://github.com/awillsousa/ccscore>

1.6 ORGANIZAÇÃO DO DOCUMENTO

Este documento é composto por 6 capítulos.

Neste capítulo foi apresentada a relevância do tema e o impacto do mesmo nos processos de avaliação textual automatizada. Foram apresentados ainda, motivação, objetivos e contribuição.

No capítulo 2 são apresentados os conceitos e definições relacionados ao problema estudado, bem como as teorias e métodos desenvolvidos até então. O objetivo desse capítulo será prover uma compreensão o mais completa possível do problema abordado neste trabalho, especialmente relacionado à avaliação automática de coesão. Também serão apresentados os conceitos de coesão e coerência textual, as teorias linguísticas nas quais o método utilizado se baseia e a forma de valoração utilizada.

No capítulo 3, se faz uma revisão da literatura acerca da avaliação de coerência e coesão textual, avaliação automática de redações para o português, em especial daquelas que consideram dimensões distintas nos seus processos de valoração e que, até este momento, figuram como o estado-da-arte nesta tarefa.

O capítulo 4 apresenta a base de textos dissertativos utilizada, seu processo de preparação e construção. Também é apresentada uma proposta de utilização do método como suporte de processos formativos e, nas últimas seções, as ferramentas desenvolvidas e aquelas adaptadas para resolver problemas específicos na obtenção do valor da coesão.

O capítulo 5 traz os experimentos executados para avaliar a qualidade dos resultados do método, bem como uma comparação entre os resultados obtidos e os valores dos avaliadores humanos. Também são feitas comparações com outros trabalhos semelhantes e, ao final do capítulo, uma discussão dos resultados obtidos nos experimentos.

Para finalizar, o capítulo 6 apresenta as conclusões obtidas e proposições de trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são apresentados os conceitos e definições necessários ao estabelecimento do problema estudado, visando também delimitar o foco da nossa abordagem. Iniciamos pela exposição da divisão histórica das pesquisas relacionadas ao processamento de linguagem em diferentes áreas e que, recentemente, tem mostrado sinais de um processo de convergência. Em seguida, apontamos as definições de coerência, as teorias e os algoritmos que serão utilizados em nossa proposta. Também são abordados de forma mais estrita o problema de avaliação automática de redações, bem como algoritmos e modelos utilizados.

2.1 PROCESSAMENTO DE LINGUAGEM NATURAL

Historicamente as pesquisas relacionadas ao processamento da linguagem falada e escrita tem sido conduzidas em áreas distintas e de forma separada, ainda que esses campos de estudo acabem por apresentar alguma sobreposição entre si: linguística em Linguística Computacional, processamento de linguagem natural em Ciência da Computação, reconhecimento de voz em Engenharia Elétrica e psicolinguística computacional em Psicologia (Jurafsky, 2000).

As áreas de conhecimento relacionadas à linguagem natural, cada uma com seus respectivos níveis de complexidade, podem ser categorizadas da seguinte maneira:

- Fonética e Fonologia - estudo dos sons da língua;
- Morfologia - estudo do significado dos componentes das palavras;
- Sintaxe - estudo das estruturas de relacionamento entre as palavras;
- Semântica - estudo do significado das construções da linguagem;
- Pragmática - estudo de como a linguagem é utilizada para atingir objetivos;
- Discurso - estudo das unidades linguísticas relacionadas a enunciados maiores e mais complexos.

Neste trabalho, cujo foco está na avaliação de coesão textual, serão abarcados elementos que vão desde a análise morfológica, sintática e semântica dos textos.

2.2 COESÃO E COERÊNCIA TEXTUAL

Nesta seção serão apresentados os fundamentos teóricos da área de linguística computacional considerados neste trabalho, bem como a demarcação da amplitude de como estes conceitos serão aqui utilizados. Assim, não será feita distinção entre texto e discurso, ainda que nos estudos de linguística moderna esses conceitos apareçam separados. Também não utilizaremos uma delimitação profunda entre os conceitos de texto e não-texto. Convém ressaltar que aqui estamos buscando identificar como a construção textual foi organizada em níveis de coesão, de forma a dar-lhes um valor de acordo com a sua qualidade, e os processos de determinação **do que pode ou não ser considerado texto é indiferente para o método aqui apresentado.**

2.2.1 Teoria do Discurso

A análise de um texto, principalmente quando se busca avaliar a sua capacidade comunicativa, é obtida por meio da análise discursiva. Ainda que esta venha a ser a abordagem mais adequada e completa, tanto da estrutura quanto da qualidade de um texto, o seu uso é dificultado pela necessidade de anotações das estruturas de discurso dentro de um texto. A anotação destas estruturas exige conhecimento linguístico acerca das áreas correlacionadas, de maneira que as informações obtidas da pragmática, da semântica e das outras áreas de estudo, sustentam as construções do discurso livre. Sendo este, o responsável pelo processo, tanto de transmissão da informação, quanto da criação de uma imagem mental do conteúdo do texto por parte do leitor, permitindo que este possa estabelecer as conexões necessárias ao entendimento de ideias. Durante esse processo de estabelecimento das conexões, os elementos mais superficiais e objetivos do texto são utilizados como guia ou como definidores destas conexões.

Considerando os objetivos primários, neste estudo serão observados três aspectos: estrutura linguística, estrutura intencional e o estado de atenção. De acordo com Nobre (2011), a estrutura linguística de um texto compreende suas divisões em unidade de discurso. A estrutura intencional abrange as intenções que estão na base comunicativa do discurso e como as relações entre as unidades de discurso ajudam a perceber essas intenções. O estado de atenção é o modelo do foco de atenção durante o discurso. A avaliação da estrutura intencional foge ao escopo e proposta deste trabalho. A análise destas estruturas é complexa principalmente por conta da ambiguidade intrínseca da linguagem, o que faz com que uma única palavra ou sentença tenha múltiplos e distintos significados, podendo sofrer alteração na ideia que expressa por conta das intenções de comunicação. Estas intenções de comunicação podem ser estabelecidas a partir das relações entre as unidades de discurso, enquanto o estado de atenção depende do modelo do foco do discurso utilizado. Dois modelos de estado de atenção podem ser utilizados para tanto: Teoria da Centragem (Grosz e Sidner, 1986) e Teoria do Foco (Sidner, 1979).

A Teoria da Centragem formula uma noção de conexão no texto, tentando explicar a motivação por trás da escolha de um discurso em detrimento de outro. Ela também busca relacionar o estado de atenção e explicar a inadequação do uso de elementos de composição das unidades básicas do discurso, por meio da observação do contexto circundante das sentenças e das entidades que nela aparecem (Nobre, 2011). A avaliação de contexto e de entidades não ocorre de maneira global, dado que o estado de atenção modelado considera apenas aspectos locais. Assim, ao avaliar uma sentença, bem como aquelas que a circundam, potenciais centros de atenção surgem. A influência desses centros potenciais é proporcional à proeminência de suas entidades, que é determinada pela função gramatical destas. Com o sujeito sendo melhor classificado que os objetos, e estes, por sua vez, melhor classificados que as demais funções. As alterações no foco do discurso refletem as transições entre diferentes potenciais centros de atenção. Assim, podemos ter, em diferentes segmentos de discurso, as seguintes situações:

1. O centro de atenção não muda e o discurso permanece centrado na mesma entidade;
2. O centro de atenção não mudou, mas provavelmente irá mudar na próxima sentença, com a sentença atual servindo como elo de ligação;
3. O foco do discurso mudou.

Por meio das possíveis transições do foco do discurso, a Teoria da Centragem se propõe a explicar a coerência dentro de um mesmo segmento de discurso, modelando a coerência local, por meio das possíveis transições dos centros de atenção entre sentenças (Nobre, 2011).

Já a Teoria do Foco, busca a partir da ideia de foco e suas estruturas de armazenamento e considerando as entidades e as suas representações semânticas, avaliar as conexões entre elementos do discurso como parágrafos e sentenças, determinando um fator de coesão global por meio do estabelecimento de focos entre pares de sentenças, bem como do conjunto de parágrafos.

A Teoria do Foco propõe um algoritmo que busca reduzir as possibilidades de antecedentes a serem utilizados em resolução de anáforas dentro do texto. Por isso recebe também o nome de Algoritmo de Focalização. Este define o foco do ator e o foco do discurso que, por sua vez, determinam o foco da sentença atual, uma lista de indivíduos distintos do foco corrente presentes na mesma sentença e uma pilha dos focos anteriores. Desta maneira o algoritmo consegue determinar a coesão existente entre as sentenças e o conjunto de parágrafos, determinando como a coesão global do texto se constrói.

2.2.2 Coesão Textual

Os elementos estruturantes de um texto que permitem que o mesmo cumpra a sua função comunicativa são:

1. coesão, que está relacionada com a conexão entre as ideias do texto;
2. coerência, que é responsável por prover sentido ao texto;
3. intencionalidade, que está relacionada com aquilo que se deseja expressar por meio do texto;
4. aceitabilidade, que é um papel desempenhado pelo leitor ao buscar entender o conteúdo do texto;
5. situacionalidade, que está relacionada ao contexto no qual o texto está situado;
6. intertextualidade, que ocorre quando elementos de outros textos são referenciados de forma direta ou indireta.

A coesão é um elemento de estruturação textual fundamental para a correta transmissão e interpretação de um texto. Sua construção se dá por meio de elementos como pronomes, advérbios, termos, expressões e até mesmo pelos sinais de pontuação (Dias, 2016) e ela se estabelece por meios dos processos de interlocução, fazendo com que o texto faça sentido para o leitor, estando também relacionada com a interpretabilidade e inteligibilidade textual (Koch, 1999). Este conceito pode ser notado em um texto por meio da avaliação da estruturas linguísticas, do sequenciamento de entidades e elementos textuais, mas não se limita apenas aos aspectos gramaticais. A coesão apresenta características semânticas, em concordância com a definição de Hallyday e Hassan (1976) que ela é a relação semântica entre os elementos do texto, no que refere às relações de sentido existentes em seu interior e que o definem. Assim, ao abordar a avaliação de coesão, pode-se utilizar os seus mecanismos de estabelecimento ou a forma como as relações de coesão se apresentam. Neste trabalho utilizou-se esta última abordagem.

Em (Koch e Travaglia, 1989) propõe-se que a coesão pode ser abordada, não pelos seus mecanismos de estabelecimento, mas sob os aspectos referencial e sequencial. Na coesão referencial, um primeiro elemento do texto (forma referencial) remete a um segundo elemento (referente textual ou antecedente). Já na coesão sequencial há uma união de elementos para dar ideia de continuidade e sequenciamento, explicitando a progressão textual por meio de relações semânticas de causa, condição, oposição, tempo, conformidade e finalidade.

A construção dos vínculos da coesão referencial é feita por meio de pronomes, numerais e outros elementos, com referências a outros termos presentes no texto. Por exemplo, no trecho “Mês passado Carlos me ligou. Ele estava de férias”, o pronome “Ele” faz uma referência direta ao elemento “Carlos” presente na frase antecedente. Os vínculos referenciais permitem uma retomada de conceitos e entidades, orientando o leitor a recuperar informações anteriores ou mantê-las ativas ou na memória de curto prazo para montar a correta associação em momento futuro. Esse processo é responsável por manter os antecedentes mais importantes, ativos durante a progressão textual, porém pode também se apresentar como um risco à qualidade textual, se utilizado em demasia ou de forma estática. Ao utilizar-se do mesmo mecanismo ou dos mesmos sintagmas de forma reiterada, não há uma gradação, nem uma proporcionalidade da saliência das entidades, conduzindo a textos com qualidade coesiva pobre, podendo transmitir a informação de forma errada (Nobre e Pellegrino, 2010).

Para um melhor entendimento, considere-se o trecho de texto “Lívia, atriz famosa, casou-se com Carlos. Carlos se diz muito feliz com a venezuelana. Carlos espera novos papéis para ele e para a famosa atriz”, onde percebe-se uma transição gradual no ganho de informações sobre “Livia”, enquanto pouco se pode inferir sobre “Carlos”, exceto que este, provavelmente, seja ator também. Contudo, o tempo de processamento dessa informação pelo leitor, é mais demorado e exige um maior esforço, além de termos um texto com uma forte estrutura coesiva, porém sem fluidez, o que faz com que o leitor tenha que tomar cada elemento do texto - sentença - buscando estabelecer as relações.

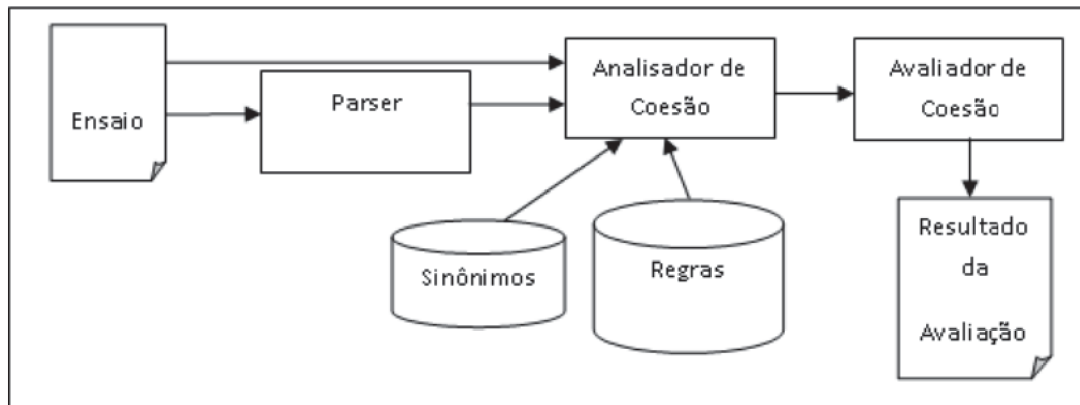
Para entender como a coesão pode ser abordada a partir do aspecto sequencial, ao considerar o trecho de texto “Lívia e Carlos, ambos atores famosos, casaram-se e parecem muito felizes juntos. Provavelmente, porque retornarão para atuar em uma nova novela”, é possível perceber que há um sequenciamento das ações, o que permite ao leitor estabelecer a ordem dos acontecimentos e as relações entre eles, bem como situá-los em relação à sua própria linha do tempo. Essas relações desempenham papel de encadear os segmentos textuais, levando à frente o discurso, sendo responsáveis por uma progressão temática, inserindo articulação e argumentação no texto (Dias, 2016). Merece destaque o fato de que alguns dos elementos coesivos responsáveis por estabelecer o processo de sequenciamento nem sempre podem ser considerados como os mais adequados em textos discursivos, em especial aqueles de processos avaliativos, nos quais o uso de expressões repetitivas ou desgastadas pelo uso recorrente devem ser evitados. Avaliação de situações desse tipo não serão alvo deste estudo, ainda que os mesmos possam vir a ser codificados por meio de atributos obtidos de textos utilizados como referencial.

Para os fins propostos neste trabalho, adotou-se a avaliação de coesão utilizada em (Nobre, 2011), pois a sua proposição se alinha aos nossos objetivos de obter informações a partir das estruturas do texto, sem necessitar de especialistas. A proposta original de Nobre (2011), segue o esquema funcional apresentado na Figura 2.1.

Como a adaptação da proposta, foi possível obter um método cuja análise das estruturas coesivas é feita por meio de informações morfossintáticas e semânticas, cuja anotação é feita por um parser. De posse das anotações parser, em conjunto com um conjunto de regras, de análise de referências e sequenciamento das relações, executa-se o processo de valoração do grau de coesão do texto. Essa abordagem, considerando o uso das diferentes estruturas do texto, que variam desde os elementos menores e mais básicos até os mais complexos e longos, permite que a valoração abarque tanto a coesão referencial quanto a sequencial.

A análise das diferentes formas da coesão e a identificação das estruturas que a suportam se dá por meio da construção de listas de focos, que se podem se apresentar como explícitos e implícitos. A lista de foco explícito é feita por meio de referências diretas a elementos do texto, como sinônimos, hipônimos e hiperônimos. Enquanto a lista de foco implícito é construída

Figura 2.1: Estruturação do processo de análise de coesão.



Fonte: (Nobre, 2011)

por meio de das anotações semânticas dos elementos do texto. Assim, elementos de estruturas maiores, como sentenças e parágrafos, que compartilhem o mesmo papel semântico, figuram na mesma de foco.

Durante o processo de formação da lista de foco explícito entre pares de sentenças ou parágrafos, se não forem encontradas relações entre os elementos, será criada uma lista intermediária formada por referências ou sinônimos dos elementos analisados que compartilhem a mesma anotação semântica. Essa lista intermediária busca avaliar se, quando não há referência explícita direta, é possível encontrar ao menos, conexões por meio dos sinônimos e dos seus papéis semânticos. Essa lista intermediária, pode ser utilizada em complemento ou substituição à lista principal para análise.

No processo de análise e avaliação de (Nobre, 2011), no qual o texto original é referenciado como ensaio, este deverá ser fornecido e de posse de uma lista de regras específicas e de uma lista de sinônimos fornecida pelo dicionário TeP2 (Maziero et al., 2008), busca-se determinar os problemas encontrados nos textos, de acordo com os seguintes passos:

1. Etiquetagem das informações morfológicas, sintáticas e semânticas por meio de um parser;
2. Execução do processo de análise da coesão, onde é feita a criação de uma lista de focos implícitos (FI) e de focos explícitos (FE) para cada par de frases e de parágrafos para, em seguida, analisar a coesão baseando-se na união e interseção dos componentes das listas de foco;
3. Cálculo do peso das relações estabelecidas por meio do módulo de avaliação;
4. Entrega do resultado com uma valoração final para a coesão textual.

Na última etapa do processo, o Algoritmos 1 é utilizado para avaliar, a coesão local. Essa avaliação é feita por meio das interseção dos focos explícito e implícito dos pares de sentenças consecutivas do texto. Assim, dado um par de sentenças (S_i e S_{i+1}), com as listas de foco explícito e implícito dadas respectivamente por FE_i , FE_{i+1} , FI_i e FI_{i+1} , com a relação entre as listas sendo calculada por $RFFE_{i,i+1} = FE_i \cap FE_{i+1}$ e $RFFI_{i,i+1} = FI_i \cap FI_{i+1}$, onde $RFFE_{i,i+1}$ e $RFFI_{i,i+1}$ indicam o relacionamento entre as listas de focos explícito e implícito de duas sentenças subsequentes.

Algoritmo 1 Avaliação da Coesão Local - CL. Fonte: (Nobre, 2011)

Require: Conjunto de todas as frases, com cada frase f identificada de forma única - F

```

for all  $f \in F$  do
    Gerar lista de Foco Explícito -  $FE$ 
    Gerar lista de Foco Implícito -  $FI$ 
    Gerar lista de Foco Explícito Intermediária
    Por meio de regras específicas, localizar elementos que afetem a coesão local
end for

for  $i \leftarrow 1, 2, \dots, \text{tamanho}(F) - 1$  do
     $f_i \leftarrow i$ -ésima frase
     $f_{i+1} \leftarrow (i + 1)$ -ésima frase
    Calcular a interseção ( $RFFE_{i,i+1}$ ) entre os elementos de  $FE_i$  e  $FE_{i+1}$ 
    Excluir os elementos em  $FE_i$  e  $FI_{i+1}$  que são representados em  $RFFE_{i,i+1}$ 
    Calcular a interseção ( $RFFI_{i,i+1}$ ) entre os elementos do  $FI_i$  e  $FI_{i+1}$ 
    Analisar  $RFFE_{i,i+1}$  e  $RFFI_{i,i+1}$  com base na Tabela 2.1.
end for

```

Para o cálculo da coesão global, o Algoritmo 2, constrói as listas de foco explícito e implícito dos parágrafos baseados nas sentenças de compõem os parágrafos.

Algoritmo 2 Avaliação da Coesão Global - CG. Fonte: (Nobre, 2011)

Require: Conjunto de todos os parágrafos, com cada parágrafo P identificado de forma única - P

```

for all  $p \in P$  do
    Criar lista de Foco Explícito dos parágrafos -  $FEP$ 
    Criar lista de Foco Implícito dos parágrafos -  $FIP$ 
end for

for all  $par(P_i, P_{i+1})$  com adjacente máxima ( $P_1eP_2, \dots, P_2eP_n, \dots, P_{n-1}eP_n$ ) do
    Calcular a interseção ( $RFEP_{i,i+1}$ ) entre os elementos do  $FEP_i$  e  $FEP_{i+1}$ 
    Excluir os elementos em  $FIP_j$  e  $FIP_{j+1}$  que são representados em  $RFEP_{j,j+1}$ 
    Calcular a interseção ( $RFIP_{j,j+1}$ ) entre os elementos do  $FIP_j$  e  $FIP_{j+1}$ 
    Analisar  $RFEP_{j,j+1}$  e  $RFIP_{j,j+1}$  com base na Tabela 2.1.
end for

Localizar elementos que geram ambiguidade e contradição em todo um parágrafo

```

As relações entre os parágrafos é obtida por meio das equações $RFEP_j = (FEP_jF_1 \cap FEP_jF_2) \cup (FEP_jF_1 \cap FEP_jF_n) \cup \dots \cup (FEP_jF_{n-1} \cap FEP_jF_n)$, para o foco explícito e a equação $RFIP_j = (FIP_jF_1 \cap FIP_jF_2) \cup (FIP_jF_1 \cap FIP_jF_n) \cup \dots \cup (FIP_jF_{n-1} \cap FIP_jF_n)$ e para o foco implícito. Assim, dado um parágrafo (P_i), deverão ser analisados todos os pares formados entre os n parágrafos do texto e aquele parágrafo específico (P_i e P_{i+1} , P_i e $P_{i+2}, \dots, P_i$

e P_n). Para cada par de parágrafos calcula-se a interseção entre as listas de foco explícito e implícito por meio das equações $RFEP_{j,j+1} = FEP_j \cap FEP_{j+1}$ e $RFIP_{j,j+1} = FIP_j \cap FIP_{j+1}$.

Tabela 2.1: Relação entre FE e FI para estabelecimento da coesão e valoração.

	$FE_i \subset FE_{i+1}$		$FE_i \not\subset FE_{i+1}$	
	Relação	Valor	Relação	Valor
$FI_i \subset FI_{i+1}$	Elaboração	1,00	Mudança de Tópico	0,50
$FI_i \not\subset FI_{i+1}$	Manutenção de tópico	0,75	Mudança de assunto	0,00

Fonte: (Nobre, 2011)

Para exemplificar o processo de análise da coesão local, na Tabela 2.2 é apresentada a avaliação e valoração da coesão entre duas sentenças, que pode ser estendida para a análise de parágrafos.

Tabela 2.2: Exemplo de avaliação da coesão para duas sentenças.

	Sentença 1	Sentença 2
	O ser humano constrói seu caráter em sociedade.	Primeiro no lar, na família e depois na Igreja ou qualquer outra instituição.
Entidades de Foco Explícito	• ser humano • caráter • sociedade	• lar • família • igreja • instituição
Entidades de Foco Implícito	• humano (H): ser humano • característica psicológica humana (f-psyq): caráter • grupo (HH): sociedade	• construção usada por humano (Lh): lar • grupo (HH): família • instituição (inst): igreja • instituição/ação (inst/act): instituição
$FE_1 \cap FE_2$	{ } (Não há entidades em comum)	
$FI_1 \cap FI_2$	"H" é o elemento em comum nas duas sentenças. Pois, "H" e "HH", podem ser entendidos como "H"	
Valor Coesão	De acordo com a Tabela 2.1, $FE_i \not\subset FE_{i+1}$ e $FI_i \subset FI_{i+1}$ resulta em Relação: Mudança de tópico, Valor: 0,50	

Fonte: Adaptado de (Nobre, 2011)

Uma vez que se tenha os valores da coesão entre os pares de sentenças (PSF) e os valores da coesão entre os pares de parágrafos com adjacência máxima (PSP), a força coesiva do texto é determinada pelo Índice de Coesão (IC), que estabelece o peso das relações entre as frases e os parágrafos, de acordo com a equação 2.1.

$$IC = \left(\frac{\sum_{i=1}^{s-1} PSF_{i,i+1}}{s-1} + \frac{\sum_{i=1}^{p-1} \sum_{j=1}^{p-1} PSP_{i,i+j}}{\sum_{j=1}^{p-1} j} \right) / 2 \quad (2.1)$$

Equação 2.1: Índice de Coesão, onde s e p indicam o total de sentenças e parágrafos do texto, $PSF_{i,i+1}$ representa a nota da relação entre frases adjacentes e $PSP_{i,i+j}$ representa a nota da relação entre os parágrafos, de acordo com a Tabela 2.1. Fonte (Nobre, 2011)

Para determinar o valor final da coesão textual, o modelo proposto por Nobre ainda utiliza dois outros índices que levam em consideração elementos que podem influir diretamente no estabelecimento da coesão e nas características do texto. O primeiro deles é o Índice de imprecisão do texto sob avaliação (IAC), que considera informações relacionadas à ambiguidades e contradições no texto. Este índice é definido pela Equação 2.2

$$IAC = (TA + TCr) / s \quad (2.2)$$

Equação 2.2: Índice de Imprecisão do texto, onde TA representa o Total de Ambiguidades, TCr representa o Total de Contradições encontradas no texto e s o total de sentenças do texto. Fonte (Nobre, 2011)

O outro índice proposta por Nobre, recebeu o nome de Índice de forma (IF) e considera problemas que afetam a coesão local, sem afetar fortemente a coesão global. Para o cálculo do índice, é feita a contagem de dos termos com problemas, como períodos incompletos, introdução de pronomes sem vínculos estabelecidos e uso demasiado de repetições como mecanismo de coesão. A quantidade de termos com problemas (TPr) é dividida pelo número de sentenças do texto (s). O IF é de calculado com a Equação 2.3.

$$IF = (TPr) / s \quad (2.3)$$

Equação 2.3: Índice de Forma do texto. Fonte (Nobre, 2011)

Na proposta de Nobre, cada um dos índices é calculado e fornecido a um modelo de inferência fuzzy para obtenção da nota relativa à coesão. Em seguida, por meio de um conjunto de regras que considera os diferentes intervalos de valores de cada um dos índices, obtêm-se um índice de coesão fuzzy. Neste estudo, não se utilizou métodos deste tipo, substituindo quando necessário a inferência fuzzy por modelos regressivos ou de aprendizagem supervisionada, uma vez que o método de inferência fuzzy necessitaria de calibrações e ajustes aos quais não se tem acesso na proposta original. Além disso, o uso de regressores tem apresentado bons resultados em trabalhos de valoração de redações e os mesmos não exigem ajustes ao mudar os dados de treinamento, podendo ser aplicados de forma direta facilitando o seu uso para distintas tarefas de avaliação de textos.

2.2.3 Coerência Textual

A coerência textual, assim como a coesão, é fator determinante na qualidade textual. Porém, ela está relacionada à representação mental do conteúdo do texto criada pelo próprio leitor.

A coerência está vinculada ao estabelecimento de um sentido para o texto, durante o processo de busca e estabelecimento de relações que o leitor realiza utilizando o seu próprio entendimento lógico, que por sua vez está relacionado a uma lógica superior denominada de conhecimento de mundo (Koch e Travaglia, 1989). Desta maneira, entende-se que a coerência abarca conceitos relacionados à inteligibilidade do texto no estabelecimento do processo de comunicação entre o leitor e o texto, considerando a capacidade do leitor de concatenar as informações textuais com seu próprio conhecimento de mundo (Nobre, 2011). É importante considerar os aspectos psicológicos do relacionamento do leitor com o texto e ressaltar que um texto coerente certamente deverá apresentar também coesão, para que a transmissão de informação ao leitor ocorra com o mínimo de ruído possível.

A relação entre coesão e coerência é entendida de forma distinta por diferentes autores. Inclusive, há autores que não fazem distinção entre coesão e coerência, utilizando os termos como equivalentes (Koch e Travaglia, 1990), como nos trabalhos encontrados na literatura, nos quais a coesão referenciada por meio dos termos "coesão local" ou "coesão microestrutural", como em (CHARROLES, 1988) e (Van Dijk et al., 1983). No estudo aqui conduzido, ainda que se apresente como uma proposta de valoração da coesão, elementos definidores de coerência também serão considerados, porém de forma implícita. Parte-se do entendimento que estes dois conceitos são complementares. De maneira que os processos de validação adotados serão baseados e comparados em valorações que levam em consideração os dois conceitos - coesão e coerência - sem contudo, abordar os aspectos de conhecimento de mundo e intenção de comunicação que, segundo Grosz (1986), fazem parte da estrutura intencional do discurso e quando considerados trariam uma abordagem mais completa da coerência.

A influência da coesão sobre a coerência pode ser percebida nos trabalhos de autores como Grosz (1986), Charroles (1988) e Koch (1990), que consideram que a coesão é a responsável por indicar ao leitor como estabelecer as conexões entre entidades, ideias e partes do discurso, o que torna a identificação dos elementos responsáveis pela coesão, e das suas relações, um passo importante no processo de avaliação de um texto. Assim, identificar as referências, conexões, agrupamentos de referências e associações com entidades do mundo real, pode prover um melhor entendimento de como o texto foi estruturado e se a sua estrutura facilita ou não a transmissão de informação e a compreensão dessa informação.

2.3 RESOLUÇÃO DE CORREFERÊNCIAS E VINCULAÇÃO DE ENTIDADES

Um importante fator no estabelecimento das relações em um texto é determinar sobre o quê ou sobre quem se está falando, bem como a quem as referências feitas dizem respeito. A determinação destas relações e das entidades que são referenciadas em um texto, podem ser feitas por meio das tarefas de PLN de menção de entidades, vinculação de entidades e resolução de correferência. Nesta seção são abordados os aspectos técnicos destas tarefas, dada a sua importância no estudo aqui apresentado.

2.3.1 Resolução de Correferências

A Resolução de Correferências é uma tarefa de PLN que consiste em identificar elementos do texto, como sentenças, termos e outros, que estão vinculadas a uma mesma referência (Fonseca et al., 2018b). A referência como mecanismo de coesão consiste em estabelecer uma relação entre um elemento do texto e uma entidade do mundo real, bem como estabelecer ligação com outros elementos textuais (Jurafsky, 2000). O uso de referências pode ocorrer por meio do uso de termos, pronomes, sinônimos, hipônimos, hiperônimos e outros componentes do texto que

permitem a retomada de uma entidade, de um trecho ou de uma ideia presente no texto. No processo de referência, os elementos que referem são denominados referentes e um grupo de referentes relacionados a uma mesma entidade, recebe o nome de cadeia de correferência.

Para um melhor entendimento dos tipos de ligações que podem estabelecidas nas cadeias de correferência, considere o trecho de texto a seguir:

(1) Para o presidente da Câmara dos Deputados, Rodrigo Maia, o projeto de autonomia do Banco Central (BC) deve avançar na Câmara só após a reforma tributária andar. Para ele, o projeto sobre a instituição presidida por Roberto Campos Neto não é urgente no curto prazo.

Os elementos *presidente da Câmara dos Deputados*, *Rodrigo Maia* e *ele* são menções ou expressões referentes ligadas à uma entidade do discurso, e neste caso, do mundo real, Rodrigo Maia, definida como referente. No caso das três menções citadas, por serem referentes a uma mesma entidade, diz-se que as mesmas são correferentes ou que correferem.

Referências como *ele* no trecho de texto (1), são consideradas como anáforas, pois fazem referência a uma entidade que já foi introduzida no texto. Enquanto, a referência presidente da Câmara dos Deputados introduz uma referência que ainda terá a entidade referente definida à frente no texto, sendo identificada como uma catáfora. Por sua vez, entidades como *Roberto Campos Neto* que aparecem uma única vez no texto, são denominadas *singleton*. A determinação de quais elementos correferem, exige a análise do tipo de referência e se as entidades e menções formam um grupo ou uma cadeia, denominada cadeia de correferências (Jurafsky e Martin, 2019).

A resolução de correferência compreende duas etapas:

- identificar as menções de entidades
- agrupar as menções em uma cadeia de correferência

Considerando o trecho de texto (1), um algoritmo de resolução de correferência deve ser capaz de identificar as seguintes cadeias:

```
{ presidente da Câmara dos Deputados, Rodrigo Maia, ele }
{ Banco Central, BC, instituição presidida por Roberto Campos Neto }
{ projeto de autonomia do Banco Central, projeto }
```

Os referentes podem ser enquadrados em 3 tipos distintos: **referentes específicos**, **referentes não-específicos** e **referentes abstratos** (Fonseca et al., 2018b). Os referentes específicos são aqueles que se utilizam, em geral, de nomes próprios ou pronomes referenciais, como aqueles mostrados no trecho de texto (1). Já os referentes não-específicos, são aqueles que se referem à uma entidade de representação coletiva ou conjunta, como as referências “Os moradores” e “Policiais” nos trechos de texto a seguir:

(2) Os moradores estavam todos aguardando. (3) Policiais faziam averiguações periódicas na região.

Já os referentes abstratos estão relacionados à entidades abstratas, como sentimentos, emoções, situações, estados e outros. Nos trechos de texto a seguir, os sintagmas “casamento” e “amor” são exemplos deste tipo de referente.

(4) O casamento foi uma grande experiência. (5) Amor, sublime amor que se foi.

Como alguns aspectos da coesão e da coerência são explicitados pelas relações e marcações linguísticas, a resolução de correferências se mostra importante dentro do processo de avaliação automática de textos, por ser responsável por estabelecer o encadeamento das referências e dos referentes, bem como a identificação destes últimos. No entanto, em algumas

situações, além da identificação das menções e do encadeamento destas existe a necessidade de determinar a qual entidade do mundo real uma determinada referência está relacionada. Também tem-se situações onde é preciso distinguir entre representações idênticas de diferentes entidades do mundo real. Essa tarefa de determinar a qual entidade do mundo real as referências estão relacionadas, é feita pela resolução ou vinculação de entidades.

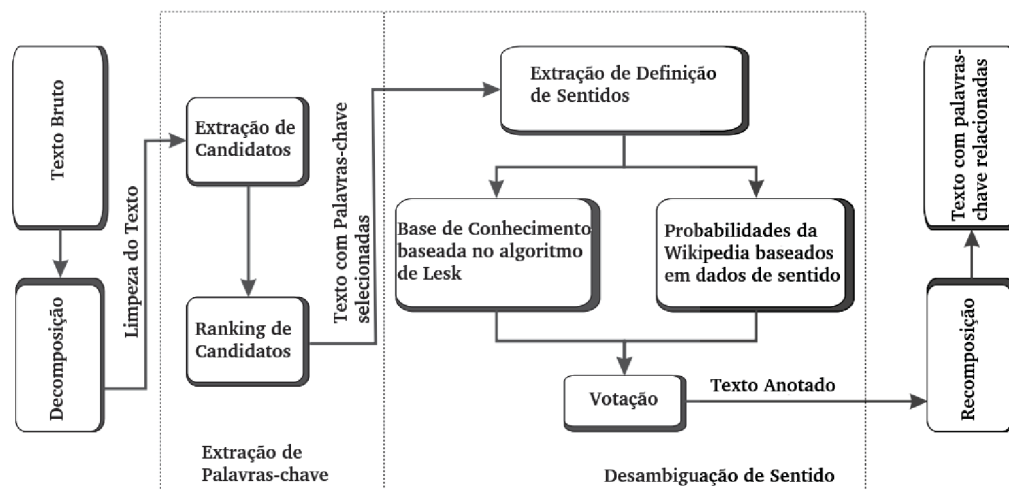
2.3.2 Detecção e Vinculação de Entidades

A tarefa de vinculação de entidades, também referida como resolução de entidades, se trata do mapeamento de uma entidade do discurso para uma entidade do mundo real (Ji e Grishman, 2011). Esse processo de mapeamento pode ser feito por meio da utilização de um gazeteer¹. Por meio de uma gazeteer se faz um mapeamento para uma ontologia, sendo a Wikipedia² a fonte mais comumente utilizada, na qual cada página serve como uma identificação única de cada entidade do mundo real (Jurafsky, 2000).

Neste trabalho, uma base de conhecimento de entidades foi criada por meio da extração de dados da Wikipedia, bem como o treinamento de um modelo de resolução de entidades.

A determinação de qual identificação ou página da Wikipedia deve ser utilizada na resolução de uma entidade é uma tarefa denominada de *wikification* (Mihalcea e Csomai, 2007). A Figura 2.2 ilustra o processo de *wikification* e suas etapas.

Figura 2.2: Processo wikification.



Fonte: (Mihalcea e Csomai, 2007)

Desde as primeiras propostas de vinculação de entidades, tem sido comum a separação da tarefa em duas etapas: detecção das menções e desambiguação das menções. Com a detecção de menções podendo ser feita por meio de algoritmos de Resolução de Entidades Nomeadas e Extração de Entidades Nomeadas. Já a desambiguação de menções pode ser feita utilizando-se métodos de aprendizagem supervisionada. Desta forma, as técnicas de vinculação e resolução de entidades podem ser utilizadas como uma forma de enriquecer os dados a serem utilizadas durante o processo de resolução de correferência, facilitando assim, os processos de avaliação e valoração de textos.

¹Um espécie de dicionário de termos, com informações adicionais

²<https://www.wikipedia.org/>

2.4 AVALIAÇÃO AUTOMÁTICA DE REDAÇÕES

A Avaliação Automática de Redações (AAR) refere-se a métodos e sistemas que permitam analisar um determinado texto dissertativo sem a interferência humana. Segundo Dikli (2006), a AAR é a tecnologia computacional que avalia e valoriza textos, com aplicações que vão desde o suporte a professores em salas de aula à correção massiva de textos em larga escala. Sistemas deste tipo são criados para auxiliar em processos de correção de exames aplicados em larga escala, mas também podem ser utilizados para processos educacionais em sala de aula (Dikli, 2006). Estes sistemas recebem um texto e partindo de critérios e regras que definem como deve ser avaliado e valorado e provêm um retorno sobre a qualidade de aspectos ou competências específicas, além de uma valoração global para o texto e valorações específicas para as competências. Todavia, são mais comuns, sistemas de AAR que devolvam simplesmente um valor, sem contudo prover retorno sobre a qualidade ou estruturação.

Diferentes sistemas de AAR estão disponíveis para a língua inglesa e alguns destes se apresentam como produtos comerciais, sendo considerados casos de sucesso no uso de métodos de avaliação automática. Como exemplo, tem-se o uso do sistema e-rater para a correção automática de testes como o TOEFL³ e o GRE⁴. Além destes, sistemas como o Criterion, Intelligent Essay e IntelliMetric também tem sido usados para tarefas de AAR. Estes sistemas demonstraram que é possível o seu uso em substituição ou em complementação ao uso de avaliadores humanos.

Mesmo em face do sucesso de sistemas de AAR para a língua inglesa, o seu uso apresenta uma forte dependência de grandes conjuntos de textos anotados de forma manual por especialistas e para cada tarefa ou categoria de textos que se deseja avaliar, é necessário criar um conjunto específico de textos anotados. Em face desta limitação, sistemas que permitam a avaliação de forma independente de um corpus anotado, podem auxiliar na execução de tarefas de avaliação e permitir a expansão em línguas diferentes do inglês, como por exemplo o português brasileiro. Ainda que nos últimos anos, alguns corpus tenham sido criados e disponibilizados publicamente, aqueles como maior volume de dados, em geral, pertencem a instituições não-públicas e não são de livre acesso. A Tabela 2.3 traz alguns dos corpus de redações em português encontrados na literatura, mas estes corpus em sua maioria, trazem apenas as valorações das competências e do texto como um todo, sem prover informações da qualidade do texto em competências específicas. Dada a limitação dos dados disponíveis, optou-se pela criação de uma base própria de redações para as avaliações a serem executadas neste trabalho, contribuindo para um aumento da quantidade de recursos deste tipo, livremente disponíveis.

A correção de textos dissertativos é feita considerando competências distintas e os valores obtidos nestas competências determinam o valor final recebido pelo texto. As abordagens onde o sistema de AAR provê apenas um valor único como forma de qualificar o texto, ou dimensões específicas do mesmo, tem recebido críticas desde as propostas dos primeiros sistemas de AAR (Dikli, 2006) e é considerada uma limitação destes sistemas. A ausência da capacidade de apontar onde, dentro do texto, estão os problemas responsáveis pelo valor calculado por um sistema de AAR, faz com esse esse tipo de sistema, provavelmente, não seja para processos educacionais. Uma vez que em processos de formação e construção do conhecimento é importante a identificação de problemas, especialmente quando se fala de construção textual.

A limitação do escopo da avaliação automática e falta de retorno quanto à qualidade textual, segundo Shermis (2013) é uma das principais falhas dos sistemas de AAR. Inclusive, ao

³O Test of English as a Foreign Language (TOEFL) é um teste padronizado que mede a proficiência na língua inglesa de falantes não-nativos.

⁴O Graduate Record Examinations (GRE) é um teste padronizado que é pré-requisito para entrar em diversas universidades nos Estados Unidos, Canadá e outros países.

Tabela 2.3: Corpus de Redações para o Português Brasileiro

Corpus	Total de Redações
Nobre (2011)*	100
Bazelato e Amorim (2013)	429
Amorim e Veloso (2017)	1840
Fonseca et al. (2018)*	56644
Bittencourt (2020)*	27184
Marinho et al. (2021)	4570

*Dados não disponíveis de forma pública

Fonte: (Autor, 2021)

avaliar os trabalhos disponíveis na literatura, percebe-se que não há uma clara distinção entre os sistemas que dão um retorno para o operador e aqueles que não o fazem. Além disso, não são considerados os aspectos que somente podem ser apreendidos por meio do processo de leitura visual do texto, algo que atualmente não pode ser feito por meios automáticos. Assim, para suporte de processos educacionais de caráter formativo, o uso de avaliação automática focada em aspectos determinados, com os métodos de avaliação automática servindo de suporte para avaliar a qualidade da escrita, parece um caminho viável. Neste estudo busca-se avaliar apenas a competência relacionada ao uso de mecanismos de coesão textual, para assim poder auxiliar educadores e estudantes nos processos de escrita por meio dessa avaliação ou da facilitação da avaliação feita por professores ou especialistas.

Para que os sistemas de AAR possam auxiliar avaliadores humanos eles devem prover retorno sobre aspectos específicos do texto (Dikli, 2006), analisando o texto a partir de diferentes dimensões e competências. Essa forma de avaliar traria uma maior explicabilidade dos resultados e da influência dos atributos na valoração final. Contudo, algumas dessas análises de dimensões específicas do texto estão muito além da capacidade dos métodos de PLN atuais, em especial aquelas relacionadas à compreensão do conteúdo do texto (Ke e Ng, 2019). Todavia, há competências que podem ser analisadas com os métodos e modelos atuais, como por exemplo a coesão, que pode ser identificada por elementos sintáticos e semânticos (Koch, 1999). Com a implementação do modelo de avaliação de coesão aqui proposto, espera-se contribuir para obtenção de melhores sistemas de AAR, tornando a sua avaliação menos superficial e a sua valoração mais explicável com relação ao uso de atributos e informações do texto.

2.5 AVALIAÇÕES E MÉTRICAS

Para avaliação dos resultados, usou-se o Kappa Quadrático Ponderado (QWK) (Cohen, 1960). O QWK é comumente utilizada em estudos de VAR, sendo uma modificação do método Kappa proposto por Cohen (1960), que mensura a diferença entre um par de valorações utilizando uma escala de 0 a 1, onde 0 indica concordância aleatória e 1 concordância total. Já o Kappa Ponderado leva em consideração as diferenças entre pares distintos de avaliações. Ou seja, alguns pares podem apresentar maior similaridade que outros e isso é considerado. Já a parte quadrática serve para penalizar as classificações erradas, o que a torna uma boa escolha para problemas de comparação de avaliação e valoração, seja entre avaliadores humanos ou entre um algoritmo e um humano, como é o caso deste estudo.

Para calcular o QWK, é necessário calcular 3 matrizes: uma matriz de pesos (W), uma matriz de escores observados (O) e uma matriz de escores esperados (E).

A matriz de pesos é obtida de acordo a fórmula:

$$W_{i,j} = \frac{(i-j)^2}{(N-1)^2} \quad (2.4)$$

Onde:

- **i** é a nota do primeiro avaliador (referência)
- **j** é a nota do segundo avaliador
- **N** é o número possível de notas que podem ser atribuídas

A matriz de escores observados é obtida de acordo com a fórmula:

$$O_{i,j} = \text{count}(i,j) \quad (2.5)$$

Onde:

- **i** é a nota do primeiro avaliador (referência)
- **j** é a nota do segundo avaliador
- **count(i,j)** é o número de textos que receberam a nota i e a nota j

A matriz de escores esperados é obtida tomando os vetores dos histogramas das notas do primeiro avaliador (u) e do segundo avaliador (v) e calculando o produto diádico⁵ ($u \oplus v$) entre eles. Em seguida, as matrizes O e E são normalizadas para que tenham a mesma soma. Assim, utilizando as matrizes, W , O e E , o valor do QWK é calculado utilizando a equação 2.6.

$$QWK = 1 - \frac{\sum_{i,j} W_{i,j} O_{i,j}}{\sum_{i,j} W_{i,j} E_{i,j}} \quad (2.6)$$

O valor fornecido pelo QWK é um número no intervalo $[-1,1]$. Quando estes valores são aplicados na comparação entre as notas dadas por um avaliador e as notas de um modelo, são interpretados da seguinte forma:

- $QWK = 0$: o modelo não é melhor que um modelo com previsões totalmente aleatórias;
- $QWK < 0$: o modelo é pior que um modelo com previsões aleatórias;
- $QWK > 0$: o modelo é melhor que um modelo com previsões aleatórias;
- $QWK = 1$: o modelo é melhor que um modelo com previsões aleatórias, com concordância máxima entre as notas do modelo e do avaliador.

Apesar de ser comumente utilizado em avaliações de desempenho de sistemas de AAR, o QWK trata os erros de forma igual, independente de sua magnitude. Por conta disso, assim como também ocorre em avaliações de VAR, como complemento, utilizou-se a Raiz do Erro Médio Quadrático (RMSE), pois esse método penaliza valores que desviem muito das notas de referência e é mais indicado para avaliar modelos regressivos.

⁵O produto diádico $u \oplus v$ é equivalente à multiplicação matricial uv^T , sendo u representado como um vetor coluna $m \times 1$ e v como um vetor coluna $n \times 1$

2.6 CONCLUSÃO

Neste capítulo apresentou-se os conceitos de coesão e coerência textual que serão abordados neste estudo, o método de avaliação da coesão local e global, bem como o fluxo de avaliação e valoração, os algoritmos, equações e valores de transição nos quais o método se baseia. Também foram apresentados os conceitos das tarefas de resolução de correferências e vinculação de entidades, que foram de grande importância no desenvolvimento do trabalho. Ao final do capítulo foram apresentadas as métricas utilizadas para comparar os resultados obtidos, conceituados os sistemas de AAR e mostrou-se como a análise da coesão pode ser utilizada nestes sistemas como uma forma de melhorá-los e de adequá-los para outros usos.

3 REVISÃO BIBLIOGRÁFICA

Neste capítulo apresenta-se uma revisão dos trabalhos relacionados à análise de coesão e coerência como forma de explicitar que o desenvolvimento dos sistemas de AAR e das teorias de análise do discurso ocorreu de forma separada, mas com esta última sofrendo mudanças de rumo a medida que a outra evoluía. Essas teorias foram criadas para prover explicabilidade sobre como as interações entre os agentes do processo de comunicação se estabelecem e quais os meios mais importantes para tal. Também são apresentadas a evolução das diferentes proposições de sistemas de AAR e algumas das lacunas apresentadas por esses sistemas, que ao longo do tempo foram alvo de críticas. Ainda busca-se mostrar que os estudos de AAR disponíveis, voltados para o português, da mesma forma que nas propostas voltadas para outros idiomas, mesmo quando consideram aspectos como a coesão de forma exclusiva, não buscam fazê-lo de forma a prover retorno de maneira explícita e detalhada das relações dos componentes textuais de alta granularidade, como sentenças e parágrafos. Inicialmente faz-se uma contextualização de como a coesão e coerência tem sido abordadas na literatura relacionada.

3.1 AVALIAÇÃO DE COESÃO E COERÊNCIA

Aos buscar na literatura estudos voltados para a análise de textos, é comum encontrar tanto estudos que fazem avaliação conjunta da coesão e coerência, quanto estudos que as avaliam em separado. Estas diferentes abordagens, possivelmente, se devem ao fato da pesquisa relacionada à AAR e à estruturação da teoria do discurso terem se iniciado em épocas próximas. Ao passo que, enquanto a pesquisa em avaliação de textos se iniciava com o trabalho de Page (1966), as teorias delineadoras da análise discursiva tiveram seu início com os trabalhos de Halliday & Hassan (1976), e a partir daí, as duas evoluíram em paralelo. Dentre as discussões relacionadas à teoria do discurso que impactaram diretamente na análise de textos, têm-se a separação entre coesão e coerência e como uma influência a outra. Tema sobre o qual, baseado na exploração da literatura aqui realizada, entende-se que a separação entre os conceitos é atualmente aceita, tanto que a avaliação de redações dissertativas as considera em separado. Considerando essa separação, optou-se neste trabalho por abordar a coesão, pois a mesma pode ser avaliada a partir de elementos do texto de forma objetiva, não necessitando de informações externas e sendo desnecessário lidar com conceitos de elevado nível de subjetividade.

Como exemplo de uma proposta que traz a definição da coesão léxica, temos o trabalho pioneiro de Halliday & Hassan na investigação da coerência de tópicos. Suas pesquisas ainda não abordavam níveis mais complexos como a pragmática, os elementos e estruturas de alto nível do discurso, cujas bases seriam estabelecidas mais tarde, com a investigação do entendimento do discurso textual (Van Dijk et al., 1983) e o estabelecimento da estruturação formal do discurso e dos modelos de organização, que teriam importante papel na avaliação da organização e na compreensão da forma como as ideias são expostas no texto. Os estudos na área de teoria do discurso que se seguiram aos trabalhos de Halliday & Hassan buscaram modelar a coerência local em domínios específicos. Provável razão pela qual estes estudos não demonstraram capacidade de transferir os conhecimentos adquiridos em um domínio para outros. A incapacidade ocorre por conta das informações externas, como a anotação das estruturas de discurso, que são específicas do tipo de texto em análise. Uma proposição recente de estudo que superou essa limitação de transferência de domínio, pode ser encontrado em (Xu et al., 2019), porém este trabalho se volta apenas para a coerência local, ignorando a coesão ou considerando que a mesma esteja implícita

no processo de avaliação da coerência, de maneira semelhante ao que ainda ocorre em estudos recentes.

Um exemplo de estudo aprofundado e voltado exclusivamente para os elementos definidores da coesão, pode ser encontrado nos trabalhos de Graesser et al. (2004), Crossley & McNamara (2010) e McNamara et al. (2010). Baseados nestes estudos, Crossley, Kyle & McNamara desenvolveram uma ferramenta denominada Tool for the Automatic Analysis of Cohesion (TAACO) (Crossley et al., 2016), que aborda um conjunto de índices que, segundo os autores, influenciam na coesão textual. Na Tabela 3.1 são apresentados os principais índices utilizados. Para validar os seus resultados, os autores fizeram comparações com avaliações de especialistas feitas em um conjunto de 313 redações escritas por calouros de cursos de graduação de uma universidade.

Tabela 3.1: Grupos de índices utilizados no TAACO.

Grupo de Índices	Descrição
Lexical overlap	Total de sobreposições de lemas para sentenças e parágrafos subsequentes
Semantic overlap	Total de sobreposições entre palavras e os conjuntos de sinônimos de acordo com a WordNet (synsets)
Givenness	Medidas de coesão que tenham reflexo na quantidade de informação que pode ser recuperada pelo leitor
Type-token ratio	Medidas da repetição de determinados tipos de palavras dividida pelo total de palavras
Connectives	Medidas baseadas nos conectivos do texto

Fonte: Adaptado de (Crossley et al., 2016)

Abordagens semelhantes ao trabalho de Crossley, Kyle & McNamara, até onde se tem conhecimento, não encontram equivalentes na língua portuguesa e a adaptação do seu método para outras línguas exige uma avaliação de todos os índices baseados em listas de elementos específicos da linguagem do estudo original. Contudo, métodos de avaliação da coesão para a língua portuguesa podem ser encontrados em trabalhos direcionados para problemas de sumarização automática extrativa e abstrativa. Como exemplo destes trabalhos, tem-se (Jorge et al., 2014), (Dias e Pardo, 2015) e (Silva, 2016). Estes trabalhos podem ser adaptados para outras tarefas de PLN, que não apenas a sumarização automática, mas apresentam a limitação de usar sistemas de regras ou modelos que abordam apenas a coerência local e não possuem mecanismos que permitam identificar o processo construtivo da coerência por meio das relações entre os elementos do texto. O mapeamento deste processo, pode ser obtido por meio de uma análise das relações coesivas, já que estas servem como um guia para criar uma representação mental coerente do texto por parte do leitor. Propostas de abordagens distintas daquelas voltadas para avaliação de textos resultantes de tarefas de sumarização podem ser encontradas em trabalhos

mais recentes, como em (Lai e Tetreault, 2018) onde os autores propõem uma avaliação do nível de coerência e da qualidade global de textos utilizando modelos neurais, modelos baseados em entidades e modelos baseados em grafos de representações vetoriais de palavras, considerando tanto coesão como coerência. Contudo, o uso destes modelos só é possível quando se dispõe de um corpus contendo anotações, o que exige conhecimento do domínio, além de um esforço considerável para avaliar e anotar os exemplares do corpus com a indicação do seu nível individual de coesão e coerência. Propostas semelhantes foram apresentadas em (Lin et al., 2011), (Li e Jurafsky, 2016) e (Nguyen e Joty, 2017). Neste último trabalho é apresentada uma proposta de avaliação de coerência local. Ainda que os trabalhos citados tenham como foco a avaliação da coesão e de coerência do texto, estes não provêm um retorno sobre como a coesão local e a coesão global são estabelecidas. Por isso, não conseguem indicar em que parte do texto a coerência ou a coesão do texto pode ser melhorada ou do quanto os elementos utilizados na avaliação influenciam no resultado final obtido.

Além dos estudos específicos de coesão e daqueles que avaliam esta de forma implícita, a análise da coesão textual também figura como parte do desenvolvimento de sistemas de AAR, que realizam o processo de valoração das redações utilizando atributos relacionados à coesão ou por meio de sistemas que a analisam como uma competência específica que compõe a nota final da redação.

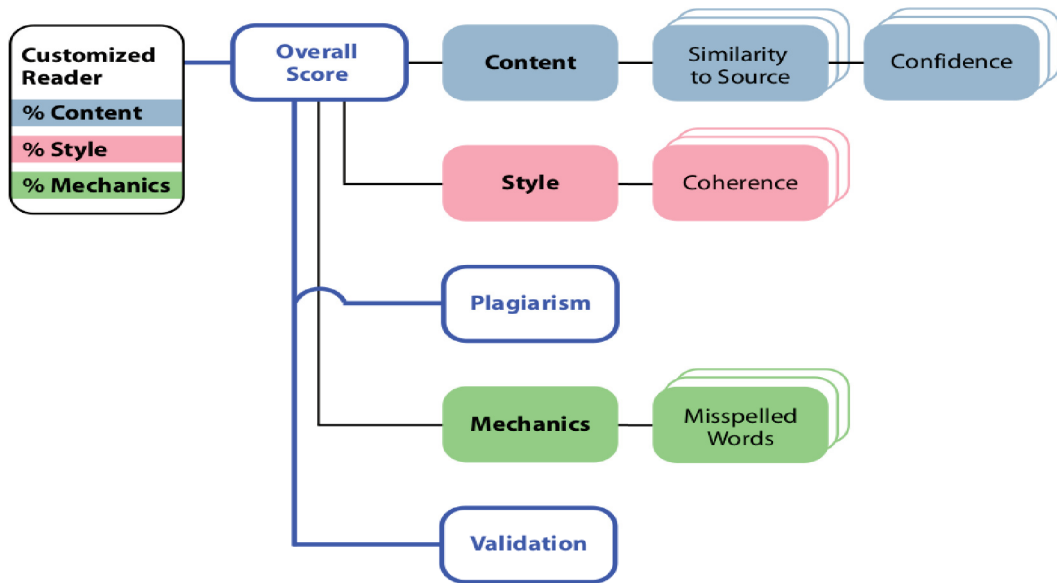
3.2 AVALIAÇÃO AUTOMÁTICA DE REDAÇÕES

O campo de estudo da AAR teve seu início com a proposta do Project Essay Grader (PEG) (Page, 1966), que é considerado como o primeiro estudo relacionado à avaliação automática de textos (Nobre e Pellegrino, 2010). Este sistema foi desenvolvido baseando-se em regras que pudessem ser implementadas à época de sua proposição, utilizando duas medidas nomeadas como *trins* e *proxes*. Os *trins* se referiam a aspectos mais superficiais do texto, como fluência, dicção, qualidade gramatical e pontuação. Enquanto os *proxes* buscavam prever a qualidade intrínseca por meio de correlações (Dikli, 2006). O sistema mostrou bons resultados para a época e continuou sua evolução, sendo utilizado ainda hoje. Apesar da sua contínua evolução e de sua valoração ser comparável a de um avaliador humano, o PEG recebeu críticas de Chung & O’Neil (1997) e Kukich (2000) por não considerar aspectos semânticos do texto e por não se aprofundar em estruturas mais complexas.

Além do PEG, outros sistemas de AAR foram propostos ao longo do tempo, com alguns destes explorando novos métodos. Um exemplar destes novos sistemas é o Intelligent Essay Assessor (IEA), que avalia o conteúdo das redações utilizando uma métrica chamada de LSA (Latent Semantic Analysis) que permite a comparação entre trechos de texto. O LSA utiliza informações de textos similares, sendo portanto, utilizado em conjuntos de textos relacionados à tópicos similares. O processo de análise do IEA considera aspectos do conteúdo, do estilo e das características mecânicas ou operacionais da escrita, como pontuação e utilização de regras gramaticais. A arquitetura do IEA está ilustrada na Figura 3.1.

Diferente de propostas anteriores, o IEA foi desenvolvido para evitar que o mesmo pudesse ser fraudado e provesse retorno sobre a qualidade do texto como um todo. Ainda que essa medida de qualidade estivesse associada à uma métrica de proximidade com o seu corpus de referência, limitando a qualidade da sua valoração à qualidade do corpus utilizado. O IEA oferecia retorno sobre aspectos genéricos da estrutura do texto - como a quantidade de palavras utilizadas - classificando-o nas categorias curto, bom e longo. Além dessa medida, o sistema também avaliava em uma escala de 0 a 5, o conteúdo e ideias, a organização, fluência das sentenças, variabilidade de palavras, uso de convenções e tipo de voz. Ao final, todas essas

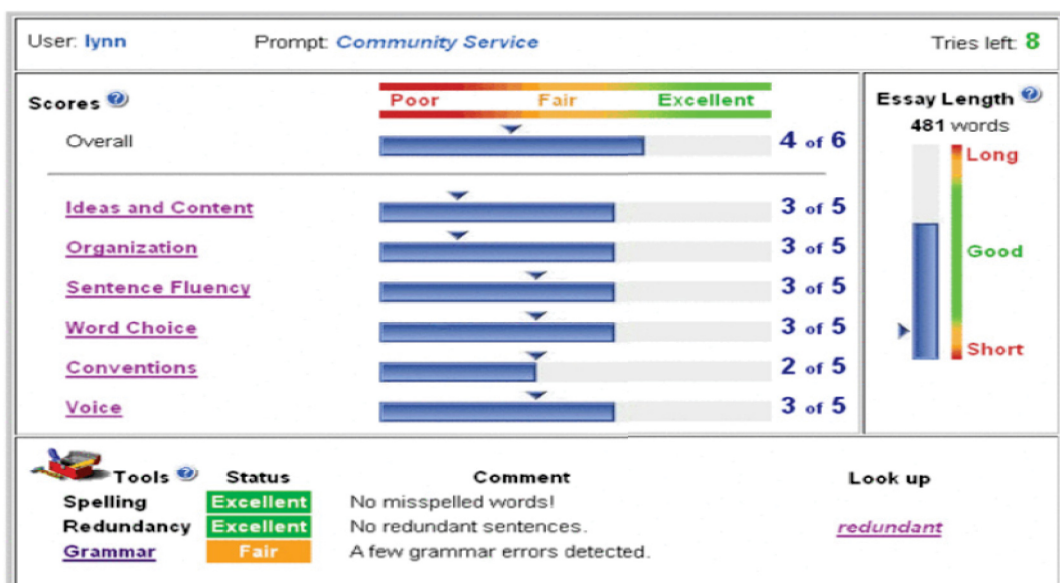
Figura 3.1: Arquitetura do Intelligent Essay Assessor.



Fonte: (Dikli, 2006)

análises compunham uma nota global, que se situava em uma escala de 0 a 6, variando entre 3 categorias distintas, que indicam que a nota final poderia ser considerada ruim, boa ou excelente. Utilizando estes mesmos termos, o sistema dava um retorno sobre a ortografia, redundância e gramática do texto. Uma tela do sistema contendo essas valorações é apresentada na Figura 3.2, onde se percebe que não há uma explicação sobre a razão dos valores obtidos ou quais foram os trechos ou elementos do texto que contribuíram em maior ou menor grau para cada nota obtida.

Figura 3.2: Tela de resultados do Intelligent Essay Assessor.



Fonte: (Dikli, 2006)

Seguiu-se à criação do PEG e do IEA diversas propostas de novos sistemas de AAR, cada uma delas utilizando propostas diferentes em consonância com a evolução da área de PLN. Todavia, todos tem em comum o fato de abordarem a valoração dos textos de forma mais global, além de exigirem um conjunto de textos semelhantes aqueles que serão avaliados. Na Tabela 3.2 são apresentados alguns dos principais sistemas de valoração automática de redações, com informações sobre as técnicas nas quais se baseiam e os aspectos do texto ao qual dão foco para obter suas valorações.

Tabela 3.2: Sistemas de valoração automática de redações.

Sistema	Desenvolvimento	Técnica	Foco Principal
PEG	Page (1966)	Estatística	Estilo
IEA	Landauer, Foltz & Laham (1997)	LSA	Conteúdo
E-rater	Time de desenvolvimento do ETS Burstein et al. (1998)	NLP	Estilo e Conteúdo
IntelliMetric	Vantage Learning Elliot et al. (1998)	NLP	Estilo e Conteúdo
BETSY	Rudner (2002)	Bayesian Text Classification	Estilo e Conteúdo

Fonte: (Autor, 2021)

Ainda que os sistemas que trabalham para obter valores mais globais da avaliação de texto tenham se mostrado equivalentes, ou até mesmo superiores aos resultados obtidos por avaliadores humanos em tarefas de correção de redações, esses sistemas não se mostram como adequados para uso em processos educacionais. Como estes sistemas avaliam textos por meio da emissão de um valor único, sem considerar aspectos específicos ou prover retorno da qualidade de competências textuais, são considerados inadequados para uso em sala de aula (Ke e Ng, 2019), onde há a necessidade de desenvolver processos de médio e longo prazo, por meio de ações formativas. Para que esses processos tenham sucesso ao serem apoiados por métodos automáticos, estes devem ser responsáveis por mostrar ao educador onde estão os problemas de escrita que merecem atenção, focando em aspectos específicos, pois dessa forma a avaliação poderia ser mais clara.

Como exemplo de propostas de escopo limitado a aspectos específicos têm-se a abordagem da coerência proposta por Higgins et al. (2004) e Somasundaran et al. (2014). Enquanto, em Louis Higgins (2010) e Persing Ng (2014), têm-se abordagens dos aspectos relacionados aos erros técnicos e aderência ao tema das redações. Um sistema que segue esta mesma linha e que merece destaque é o sistema Criterion de Burstein et al. (2004), que aborda mais de um aspecto do texto. Ao que se seguiram propostas de AAR que buscaram abordar a avaliação da qualidade de um texto a partir de um conjunto de dimensões textuais. Uma lista destas dimensões é apresentada na Tabela 3.3. Apesar de muitas destas dimensões possibilitarem análises mais profundas da qualidade do texto, a modelagem de algumas dessas dimensões exigiria uma compreensão do conteúdo do texto e isso está muito distante da capacidade dos melhores sistemas disponíveis atualmente (Ke e Ng, 2019).

Os sistemas voltados para a valoração de dimensões específicas de redações, assim como aqueles que utilizam abordagens mais holísticas, são sustentados em grande parte por processo de engenharia de atributos. Como muitos desses sistemas são baseados em métodos de aprendizagem supervisionada ou necessitam de um conjunto de dados anotados, a extração de

Tabela 3.3: Dimensões de análise da qualidade textual.

Dimensão	Descrição
Gramaticalidade	Aspectos gramaticais
Uso	Uso de preposições, utilização de palavras
Mecânica	Ortografia, pontuação, capitalização
Estilo	Escolha de palavras, variedade de estrutura de frases
Relevância	Relevância do conteúdo para o tema
Organização	O quão bem o texto está estruturado
Desenvolvimento	Desenvolvimento de ideias com exemplos
Coesão	Uso adequado de frases de transição
Coerência	Transições apropriadas entre ideias
Clareza	Inteligibilidade da tese apresentada
Persuasão	Nível de convencimento do argumento principal

Fonte: (Ke e Ng, 2019)

atributos tem um importante papel para o sucesso de sistemas deste tipo. Dentre os diversos tipos de atributos utilizados no processo de valoração, podemos destacar aqueles:

- baseados em tamanho: descrevem o tamanho de um texto a partir de suas sentenças, palavras e caracteres;
- léxicos: unigramas, bigramas e trigramas que aparecem no texto, bem como estatísticas e contagens baseadas nestes;
- baseados em embeddings: representações vetoriais de sentenças ou palavras do texto;
- de categorias de palavras: listas de palavras ou dicionários;
- de relevância do tema: codificam a relevância do texto em relação ao tema;
- de apreensibilidade: codificam o quanto um texto é difícil de ler;
- sintáticos: codificam as informações sintáticas;
- de argumentação: baseados na estrutura argumentativa do texto;
- semânticos: codificam as relações semânticas entre as palavras do texto;
- de discurso: codificam a estrutura de discurso de um texto.

Diferentes abordagens neurais e de modelagem de documentos foram propostas para sistemas de AAR, conforme mostra a Tabela 3.4. Estas abordagens apresentam potencial para obter uma modelagem mais complexa da qualidade do texto, mas a quantidade e a qualidade dos dados necessários para os seus processos podem se tornar uma barreira para o seu uso, pela escassez de recursos deste tipo como ocorre com modelos para o português.

Os sistemas de VAR têm experimentado avanços nos últimos tempos, com abordagens baseadas em arquiteturas de redes neurais. Ainda assim, o retorno sobre quais aspectos merecem mais atenção para uma melhora do texto, permanece ausente na maioria destes sistemas (Ke e Ng, 2019). Se considerarmos os sistemas propostos para o português nota-se uma escassez de propostas de valoração como um todo, sendo esta escassez maior para sistemas que avaliem as competências em separado.

Tabela 3.4: Abordagens de redes neurais e modelagem de documentos.

Abordagem	Autor	Técnica
Neural	Taghipour and Ng (2016)	Redes LSTM
Aprendizagem de Vetor de Palavras de Valor Específico	Alikaniotis et al. (2016)	Embeddings
Modelagem da Estrutura do Documento	Dong and Zhang (2016)	CNN
Pooling baseado em Atenção	Dong et al. (2017)	Mecanismos de Atenção
Modelagem da Coerência	Tay et al. (2018)	LSTM
Transferência de Aprendizagem	Daume III (2007)	Expansão do Espaço de Atributos
	Phandi et al. (2015)	Bayesian Linear Ridge Regression
	Cummins et al. (2016)	Expansão do Espaço de Atributos
	Jin et al. (2018)	Modelo Regressivo

Fonte: Adaptado de (Ke e Ng, 2019)

3.3 VALORAÇÃO DE REDAÇÕES EM PORTUGUÊS

Como um dos primeiros trabalhos - senão o primeiro - voltado para valoração de redações em língua portuguesa que leva em consideração diferentes dimensões para a valoração do texto, o trabalho de Amorim & Veloso (2017) é utilizado como referência em outros estudos semelhantes. A sua abordagem utilizou como base um conjunto de redações do ENEM, sobre o qual executou-se um processo de avaliação em relação às cinco competências exigidas no exame, por meio da extração de atributos e utilização de métodos supervisionados.

Os atributos utilizados por Amorim & Veloso foram agrupados em dois conjuntos, separados em atributos de domínio e de caráter geral, conforme apresentado na Tabela 3.5. Os atributos de caráter geral ainda foram divididos em subgrupos de gramática e estilo, atributos sintáticos, complexidade léxica e uso específico de acordo com o tema. De posse dos atributos, foram realizados dois experimentos que avaliaram o desempenho na predição de cada competência das redações e a influência que cada atributo teria na tarefa de predição da nota. Esses experimentos tiveram como referências as notas atribuídas por corretores humanos, que foram comparadas com as notas de predição do método. Para avaliar o desempenho do modelo foi utilizada a métrica do Kappa Quadrático Ponderado (QWK). As predições dos valores das notas foram obtidas por meio de algoritmos de regressão, seguindo uma abordagem que pode ser apontada como uma tendência para a época, mas que continua a dar bons resultados, mesmo hoje.

Ainda que a proposta de Amorim & Veloso seja um dos primeiros trabalhos voltados para as diferentes competências exigidas de textos dissertativos, a proposição de avaliar diferentes dimensões já estava presente no desenvolvimento da ferramenta SciPo (Scientific Portuguese) (Feltrim et al., 2006) que tinha como objetivo apoiar autores iniciantes na escrita de textos científicos, no tocante à introdução e resumo. Este sistema apresentava vários recursos, mas destaca-se o seu módulo de análise de coerência (MAC), responsável pela detecção de problemas,

Tabela 3.5: Grupos de atributos utilizados.

Grupo	Atributo
Domínio	Qtd. de ocorrências de primeira pessoa do singular de verbos e pronomes
	Razão entre a qtd. de ocorrências de primeira pessoa do singular de verbos e pronomes e o nº de tokens
	Qtd. de ocorrências de pronomes demonstrativos
	Razão entre a qtd. de ocorrências de pronomes demonstrativos e o nº de tokens
	Qtd. de ocorrências de enclise
	Razão entre a qtd. de ocorrências de enclise e o nº de tokens
Geral	Qtd. de sentenças maiores que 70 caracteres
	Qtd. de erros gramaticais
	Razão entre a qtd. de erros gramaticais e o nº de tokens
	Qtd. de erros ortográficos
	Razão entre a quantidade de erros ortográficos e o nº de tokens
	Razão entre a qtd de erros de estilo e nº de sentenças
	Qtd. de marcadores de discurso
	Razão entre a qtd. de marcadores de discurso e o nº de sentenças
	Flesh Score
	Médio do tamanho das palavras em sílabas
	Nº de tokens
	Similaridade com o tema
	Qtd. de palavras diferentes

Fonte: (Amorim e Veloso, 2017)

por meio do uso de Análise de Semântica Latente (LSA)¹. De acordo com Silva (2016) este método foi utilizada para avaliar três relacionamentos semânticos ou dimensões:

- Dimensão Título: avalia o relacionamento entre o título do resumo e o componente Propósito;
- Dimensão Propósito: avalia o relacionamento entre o componente Propósito e os componentes Metodologia, Resultado e Conclusão;
- Dimensão Lacuna-Contexto: avalia o relacionamento entre os componentes Lacuna e Contexto.

A proposta e implementação do MAC foi feita muito tempo após a proposta do SciPo, dada a contínua evolução da ferramenta. As três dimensões avaliadas pelo MAC foram propostas em (de Souza e Feltrim, 2013) e (Souza, 2011). Ainda foi feita a proposição de uma quarta dimensão de análise por Souza (2011). Essa dimensão seria relacionada às quebras de linearidade, onde o objetivo principal seria detectar problemas de coerência por meio da detecção de problema na identificação da ligação entre as sentenças. Todavia, os resultados obtidos não foram satisfatórios e o autor sugere que o uso de um modelo baseado em entidades, como o modelo de grade de entidades de Barzilay e Lapata, poderia ser mais adequado; sugestão esta desenvolvida no trabalho de Freitas (2013), que demonstrou que, por meio de corpus de resumos científicos

¹do inglês Latent Semantic Analysis

anotados manualmente, seria possível identificar as quebras de linearidade. Esta análise, apesar de prover um retorno sobre as possíveis quebras de linearidade, não provê onde no texto há problemas de escrita que mereçam atenção do autor, limitando-se ao resumo.

Assim como a proposta de uso do modelo de grade de entidades, as propostas de abordagem da coerência e coesão por meio da Teoria da Centragem e da Teoria do Foco, tem nas entidades do texto e seus papéis sintáticos e semânticos as principais informações para determinar como os elementos textuais estão conectados. Contudo, diferentemente do modelo de grade de entidades, a descoberta dos vínculos estabelecidos nos processos coesivos do texto podem ser feitos baseando-se apenas em suas informações semânticas e sintáticas, bem como por suas referências anafóricas e cadeias de correferência. Abordagens deste tipo são propostas em (Nobre e Pellegrino, 2010) e (Nobre, 2011). Nestes dois trabalhos os autores determinaram um método para valorar a qualidade dos textos produzidos, sendo possível comparar textos distintos, a partir da valoração da coesão e coerência local e global dos textos.

Nobre (2011) propôs um sistema de AAR baseado na Teoria do Foco e Teoria da Centragem, apoiado por regras específicas para buscar problemas em redações dissertativas, avaliando quesitos como a coesão, a coerência, adequação ao tema e estruturação do texto. Os valores obtidos para os diferentes quesitos abordados foram apresentados a um modelo de decisão baseado em lógica fuzzy. Utilizando como métrica a raiz do erro médio quadrático (RMSE), o autor reportou resultados equiparáveis aos de um avaliador humano. Inclusive, na proporção da divergência entre as notas atribuídas pelo seu sistemas de VAR com as notas atribuídas pelos avaliadores humanos. O método de Nobre não fez valorações sobre competências específicas - ainda que o seu método disponha de capacidade para tal - nem tão pouco fornecia sugestões de melhorias ou pontos problemáticos que poderiam ser abordados. Além disso, a forma de valoração era feita por meio de uma nota final, ainda que aspectos como coesão e coerência fossem analisados em separado, para ao final, por meio de uma fórmula específica, agregar todos os valores obtidos em uma nota global.

Abordagens baseadas em aspectos semânticos ou até mesmo em teorias linguísticas são incomuns, uma vez que o processo de extração de atributos e, mais recentemente, a utilização de arquiteturas de redes neurais, tem demonstrado potencial para evoluir os sistemas de AAR. Contudo, especificamente no caso do português, métodos de regressão associados com extração de atributos específicos ainda têm demonstrado resultados melhores que arquiteturas de redes neurais (Fonseca et al., 2018a). Com o uso de redes neurais que eliminem as etapas de extração de atributos seria possível a utilização destas arquiteturas de maneira mais rápida e sem a necessidade de anotação especializada, ainda que esses modelos necessitem de um conjunto de representações vetoriais semânticas de qualidade. Em (Bittencourt Júnior et al., 2020), o autor avaliou um conjunto de 14 modelos distintos de redes neurais profundas, reportando resultados com diferença média inferior a 100 pontos para a correção automática de redações do ENEM. Valor esse que, segundo o autor, é considerado como adequado quando aplicado à diferença entre as notas de avaliadores humanos. O autor ainda reportou os resultados do RMSE de 100,496 e do QWK de 0,462 no melhor dos casos, o que indica - no caso do QWK - que os valores obtidos pelo melhor modelo não se deve a fatores aleatórios e que apresentam consistência. Contudo, não foram reportadas comparações com outros métodos ou outros estudos correlatos.

Uma outra abordagem utilizando redes neurais pode ser encontrada em (Fonseca et al., 2018a), onde os autores utilizaram tanto o processo de extração de atributos, quanto uma abordagem baseada em duas redes neurais sequenciais para gerar uma representação vetorial das sentenças e do texto dissertativo como um todo. Utilizando-se dessas representações vetoriais, foi feito um processo de aprendizagem supervisionada para determinar notas para cada uma das competências de um conjunto de redações do ENEM, bem como de uma nota geral. Os

autores reportaram bons resultados em comparação com os valores das notas atribuídas pelos corretores humanos e também, quando comparados com o trabalho de Amorim Veloso, utilizado como referência. Contudo, os melhores resultados deste trabalho foram obtidos com o uso de modelos regressivos, utilizados com um conjunto de atributos específicos contendo mais de 680 elementos. Todavia, os autores reportaram que há uma dificuldade em obter um único modelo que consiga valorar competências distintas, sendo comum que um modelo que avalie bem uma competência tenha um desempenho ruim em outra, sugerindo que uma abordagem em separado venha a ser uma possibilidade de melhoria.

3.4 DISCUSSÃO

A revisão bibliográfica foi conduzida, considerando uma janela temporal do ano de 2006 até 2020, para analisar a existência de trabalhos voltados para a avaliação de coesão textual em português, buscando trabalhos que pudessem ser aplicados à avaliação desta e que permitissem um retorno sobre a qualidade do texto ou de seus elementos coesivos. Do estudo conduzido percebe-se que para a língua inglesa já existe sistemas de AAR que permitem indicar os quesitos a serem abordados para se obter uma melhora da qualidade do texto. Contudo, para o português brasileiro há uma escassez de trabalhos relacionados, especialmente quando se busca métodos que contribuam para processos formativos, sendo mais comuns abordagens de avaliações holísticas que buscam fornecer uma valoração global e não específica por competência.

Além das questões relacionadas à abordagem e ao foco da valoração, nota-se que há muitos trabalhos que necessitam de dados anotados ou corpus específico e de alta complexidade, como aqueles contendo anotações relacionadas à teoria do discurso. O que torna a evolução da área e de novas propostas de pesquisa mais restritas, dando-se por esquecido que as teorias linguísticas possuem um valor intrínseco que merece ser explorado, mesmo à luz de métodos computacionais mais modernos como redes neurais convolucionais e sequenciais. Até mesmo pelo fato dessas arquiteturas não terem alcançado resultados muito superiores a modelos de menor complexidade como algoritmos de regressão.

Assim, a avaliação dos trabalhos relacionados mostra a necessidade de estudos que busquem métodos mais rápidos de avaliar a qualidade do texto, provendo um retorno sobre a sua qualidade e prováveis problemas responsáveis por isso. Também mostra que apesar da quantidade de trabalhos voltados para corpus anotado, a quantidade destes tipos de corpus disponíveis é baixa. Na mesma medida, também é pequena a quantidade de bases textuais abertas e de livre uso, que permitam avaliar a qualidade de textos; o que demonstra que a criação destes artefatos e outros que permitam estruturar os processos de avaliação e valoração textual é, antes de tudo, uma necessidade e também um requisito de grande importância para a evolução e expansão da área de PLN voltada para a nossa língua materna.

3.5 CONCLUSÃO

Neste capítulo apresentou-se alguns dos trabalhos encontrados na literatura relacionada, em especial aqueles direcionados para a língua portuguesa, com foco na avaliação de coesão e coerência na valoração de redações. Também foram apresentados os principais trabalhos voltados para o português que executam os processos de valoração por meio da avaliação de distintas dimensões e que em sua maioria são condizentes com as competências exigidas na elaboração de textos dissertativos - requerimento comum em exames aplicados em larga escala como o ENEM. Discutimos, à luz dos recentes avanços de sistemas de AAR, a carência para a língua portuguesa deste tipo de sistema e de artefatos que permitam construí-los.

O trabalho de investigação da literatura relacionada, aqui conduzido, foi essencial para modelar uma proposta que permita valorar os textos observando um aspecto específico, que seja possível de mensurar a partir de elementos objetivos do texto e de sua estrutura, o que faz com que o método de avaliação utilizado não dependa de informações externas ao próprio texto. Por conta disso, escolheu-se avaliar a coesão textual pois, além de apresentar estas características, oferece suporte para análise de aspectos mais subjetivos do texto, como a coerência e o discurso livre. Como a coesão é responsável pela conexão de ideias dentro do texto e os trabalhos encontrados na literatura não dão enfoque em prover um retorno acerca desta, ou de quaisquer outras dimensões específicas, percebe-se a necessidade de propostas que possam permitir este tipo de valoração e, ao mesmo tempo, dar suporte aos avaliadores.

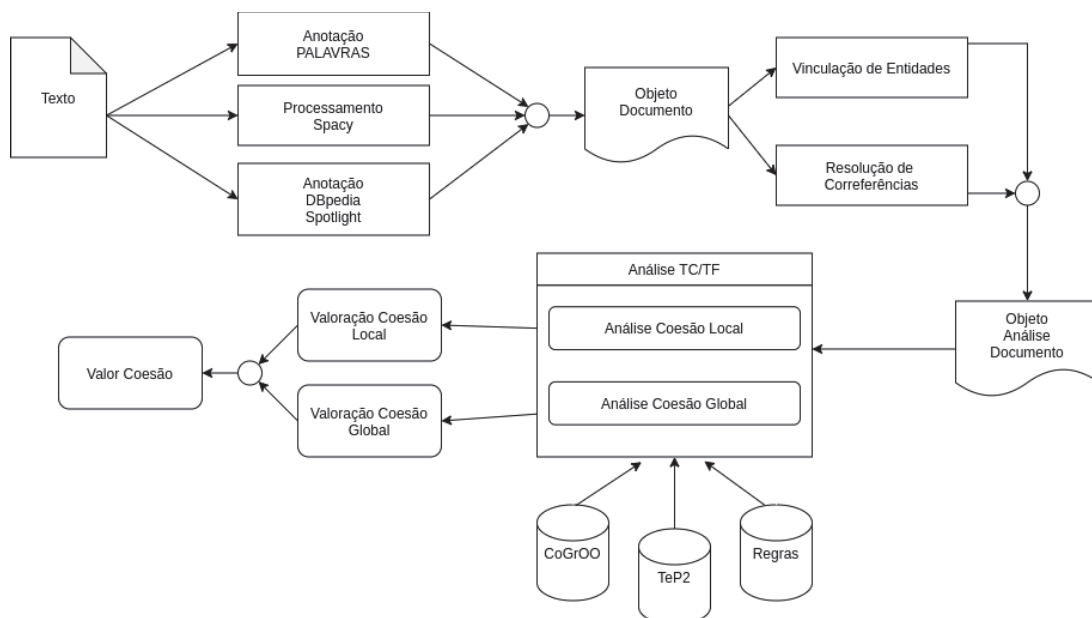
4 VALORAÇÃO E AVALIAÇÃO AUTOMÁTICA DA COESÃO

Neste capítulo apresenta-se a arquitetura proposta do método de avaliação e valoração da coesão. O método aqui desenvolvido é uma integração e adaptação de diferentes propostas existentes para tarefas como avaliação e valoração da coesão, menção e reconhecimento de identidades, resolução de correferência e correção gramatical. Destaca-se algumas soluções que, ou tiveram modelos treinados de forma específica para o português, ou tiveram que passar por processos adaptativos. Ao final do capítulo são apresentadas as comparações que serão feitas com outros trabalhos relacionados como forma de validar os resultados obtidos.

4.1 MÉTODO DE VALORAÇÃO DA COESÃO

Para que se tenha um processo de avaliação automática da coesão são necessárias diferentes ferramentas, que serão responsáveis pela detecção, vinculação e diferenciação dos componentes de cada um dos mecanismos, utilizando-se de informações sintáticas e semânticas do texto. Na Figura 4.1 é apresentada a arquitetura proposta para análise e valoração da coesão. Os textos apresentados na arquitetura fazem parte da base de redações que foi construída como parte deste trabalho e todas as etapas de preparação e pré-processamento do texto são feitas antes destes textos serem fornecidos para avaliação.

Figura 4.1: Avaliação e valoração de coesão.



Fonte: (Autor, 2021)

Ao ser avaliado pelo método aqui apresentado, um objeto de texto, no caso uma redação, passa pelas seguintes etapas:

1. Anotação de dados

- Anotação das informações sintáticas e gramaticais pelo parser do Spacy (Honnibal e Montani, 2017)

- Reconhecimento de Entidades Nomeadas pelo parser do Spacy
 - Anotação das informações semânticas pelo parser PALAVRAS (Bick, 2014)
 - Detecção das menções de entidades usando o DBPedia Spotlight (Mendes et al., 2011)
2. Agrupamento e alinhamento das anotações em relação ao texto original gerando um Objeto Documento
 3. Processamento de Entidades
 - Execução do processo de vinculação de entidades
 - Execução do processo de resolução de correferências utilizando a ferramenta CORP (Fonseca et al., 2018b).
 4. Agregação de resultados e exclusão de informações redundantes em um Objeto Análise Documento
 5. Análise baseada na Teoria do Centragem e na Teoria do Foco
 - Busca de sinônimos das entidades, menções e referências na base de sinônimos do TeP2
 - Busca dos erros gramaticais e ortográficos por meio do corretor CoGrOO (Kinoshita et al., 2007)
 - Aplicação das regras de uso das informações obtidas para decidir informações a serem mantidas e excluídas¹
 - Análise da coesão local pelo método de Nobre (2011)
 - Análise da coesão global pelo método de Nobre (2011)
 6. Cálculo dos valores
 - Cálculo dos valores da coesão local
 - Cálculo dos valores da coesão global
 7. Cálculo do valor único para a coesão

Para cada componente da arquitetura apresentada, diferentes propostas encontradas na literatura foram avaliadas e testadas para a execução de tarefas anteriores à valoração final da coesão. As escolhas foram motivadas principalmente pelos resultados reportados, disponibilidade e uso livre, visando um método eficaz, mas com reprodutibilidade viável e acessível. Para validar o método e arquitetura proposta, foi desenvolvido um protótipo em linguagem Python², utilizando bibliotecas como Scikit-learn³, Spacy⁴, NLTK⁵ e Pandas⁶ para as tarefas mais genéricas do desenvolvimento. Para as tarefas de maior escopo e complexidade foram utilizadas ferramentas específicas descritas nas seções a seguir.

¹No Apêndice A são apresentadas algumas das regras de separação dos resultados fornecidos pelo CoGroo. As demais regras dizem respeito à eliminação de duplicidades na detecção de entidades, onde detecções repetidas ou partes de entidades são excluídas do processo.

²<https://www.python.org>

³<https://scikit-learn.org>

⁴<https://spacy.io>

⁵<https://www.nltk.org>

⁶<https://pandas.pydata.org/>

4.2 ETIQUETAGEM SEMÂNTICA E SINTÁTICA

A anotação das partes do discurso (POS⁷), foi feita utilizando a biblioteca Spacy (Honnibal e Montani, 2017), conforme exemplo na Tabela 4.1, que utiliza como fontes um treebank que segue a padronização Universal Dependencies (UD) para o português e o WikiNER, uma abordagem para resolução de entidades nomeadas (REN) multilínguas que se utiliza de entidades e trechos da Wikipedia para o treinamento de modelos de REN. As anotações feitas utilizando a biblioteca Spacy, são utilizadas para identificar as classes gramaticais, por meio das etiquetas de POS. A biblioteca Spacy também foi utilizada para a resolução de entidades nomeadas.

Tabela 4.1: Anotação de texto usando a biblioteca Spacy.

Token	Lemma	Tags de POS/Dependências
Estes	este	PRON nsubj
são	ser	AUX cop
fatores	fator	NOUN ROOT
que	que	PRON nsubj
fazem	fazer	VERB acl:relcl
parte	partir	NOUN obj
da	da	DET case
atualidade	atualidade	NOUN nmod
brasileira	brasileiro	ADJ amod
.	.	PUNCT punct

Fonte: (Autor, 2021)

As anotações de etiquetas semânticas foram feitas utilizando o parser PALAVRAS (Bick, 2014), que também possui anotação de entidades, mas optou-se por não utilizá-la por conta do parser não ter uma licença de uso livre. Contudo, foi a única ferramenta que encontramos que permitiu uma anotação eficiente das etiquetas semânticas. A Tabela 4.2 traz uma amostra de um texto anotado pelo parser.

Ainda que se tenha tentado avaliar soluções que se utilizam da WordNet⁸ e do desenvolvimento de uma consulta automatizada à Wikidata^{9,10}, a dificuldade em definir um escopo e grau de profundidade da busca dos tópicos aos quais uma entidade está relacionada, não se mostrou viável.

Para detecção das menções à entidades foi utilizado um processo que utiliza a DBpedia¹¹ como base de conhecimento. A DBpedia é um projeto aberto que realiza a estruturação de dados da Wikipedia, contendo uma base de conhecimento de mais de 2,6 milhões de entidades em 30 línguas diferentes e mais de 270 milhões de triplas RDF (Bizer et al., 2009)¹².

De forma a permitir uma identificação dinâmica de entidades, Mendes et al. (2011) implementaram um algoritmo que realiza as etapas de extração de candidatos e desambiguação,

⁷Do inglês part-of-speech

⁸<https://wordnet.princeton.edu/>

⁹www.wikidata.org

¹⁰A Wikidata é uma base de conhecimento acessível de forma livre que contém dados estruturados da Wikipedia, podendo ser acessado por humanos e por máquinas.

¹¹<https://www.dbpedia.org>

¹²Um tripla RDF é uma entidade atômica de dados definida no modelo de dados do Resource Description Framework (RDF).

Tabela 4.2: Anotação de texto usando o parser PALAVRAS.

Token	Lemma	Tags Semânticas	Tags de POS/Dependências
Estes	este		DET @SUBJ>
são	ser		V PR VFIN @FS-STA
fatores	fator	<ac><Hprof>	N @<SC
que	que		SPEC @SUBJ>
fazem	fazer		V PR VFIN @FS-N<
parte	parte	<part><amount><L> <HH><sem-l><ac>	N @<ACC
de	de		PRP @N<
a	a		DET @>N
atualidade	atualidade	<am><temp><sem-l>	N @P<
brasileira	brasileiro	<nat><jh>	ADJ @N<
.			

Fonte: (Autor, 2021)

valendo-se da consulta aberta à DBpedia, anotando essas entidades em um documento de texto. Durante as etapas de anotação o documento tem suas menções marcadas de forma dinâmica utilizando os identificadores únicos - URI¹³. O algoritmo implementado é disponibilizado por meio da ferramenta DBpedia Spotlight (Mendes et al., 2011). Na Tabela 4.3 é possível ver um exemplo de um trecho de texto e uma anotação feita pela ferramenta.

Tabela 4.3: Anotação de uma entidade usando DBpedia Spotlight.

Texto	...suavizar as flutuações do nível de atividade econômica e fomentar o pleno emprego no País	
Dados DBpedia Spotlight	URI	http://pt.dbpedia.org/resource/Pleno_emprego
	surfaceForm	pleno emprego
	similarityScore	0.99

Fonte: (Autor, 2021)

A utilização da ferramenta DBpedia Spotlight, como complemento ao uso do Spacy, se mostrou de grande utilidade principalmente na detecção de entidades que seriam difíceis de serem definidas como tal pelos modelos de reconhecimento de entidades nomeadas disponíveis. Se encontrarmos dentro de um texto a ser analisado um termo como, por exemplo "pleno emprego", os sistemas de reconhecimento de entidades que foram utilizados neste trabalho, não reconhecem este termo como uma entidade. Contudo, "pleno emprego" pode ser considerado uma entidade específica em áreas como a Economia e Administração Pública. Ao utilizar a detecção de entidades da DBpedia, por conta da amplitude da categorização que a Wikipedia dispõe, o termo é corretamente identificado como uma entidade. A correta identificação do termo como entidade, influencia na construção das listas de foco implícito e explícito, que são de grande importância

¹³Uniform Resource Identifier é uma cadeia de caracteres compacta usada para identificar ou denominar um recurso na Internet.

na identificação de relações entre os elementos do texto e, consequentemente, na valoração da coesão.

4.3 DICIONÁRIOS E BASES DE SINÔNIMOS

Para a busca de sinônimos, foi utilizada base de sinônimos TeP2 (MAZIERO, PARDO, et al., 2008). Já para a obtenção das listas de hiperônimos e hipônimos, foi avaliada a base disponibilizada pelo Projeto OntoPT¹⁴. Contudo, por conta dos resultados obtidos e a dificuldade da delimitação do nível das consultas, optou-se pela utilização de mecanismos de vinculação de entidades para determinar esses relacionamentos coesivos. O trecho de texto a seguir traz um exemplo da sentença original e da lista de foco intermediário criada a partir dos sinônimos da lista de foco explícito da sentença original.

Texto Original: Houve progressos e retrocessos ao longo da história nesse âmbito.

Lista de Foco Explícito: ['longo', 'retrocessos', 'âmbito', 'progressos', 'história']

Lista Intermediária de FE: {'longo': {'copioso', 'comprido', 'extenso', 'demorado', 'largo', 'dilatado', 'estirado', 'prolixo', 'grande'}, 'âmbito': {'esfera', 'área', 'ambiente', 'atmosfera', 'circunferência', 'contorno', 'circuito', 'domínio', 'campo', 'periferia'}, 'história': {'lampana', 'conto', 'peta', 'tradição', 'tretas', 'prego', 'potoca', 'narrativa', 'mentira', 'narração', 'conversa', 'pulha', 'patranha', 'lorota', 'romance'}}

Para busca de palavras incompletas, com grafia incorreta ou que não façam parte do vocabulário do português, foram avaliados os dicionários DELA¹⁵ (DELAS, DELAF e DELACF) que trazem uma lista de palavras simples, compostas e as suas respectivas formas flexionadas. Contudo, a utilização destes dicionários em conjunto com a ferramenta de correção gramatical escolhida apresentou algumas dificuldades, então optou-se pelo uso apenas das avaliações do corretor gramatical escolhido.

4.4 CORRETOR GRAMATICAL

Para a busca de erros ortográficos e gramaticais foi escolhido o corretor CoGrOO (Kinoshita et al., 2007; Silva e Finger, 2013). Dentre os erros gramaticais detectados pelo corretor, merecem destaque:

- colocação pronominal;
- concordância nominal;
- concordância entre sujeito e verbo;
- concordância verbal;
- uso de crase;
- regência nominal;
- regência verbal;

¹⁴<http://ontopt.dei.uc.pt>

¹⁵<http://www.nilc.icmc.usp.br/nilc/projects/unitex-pb/web/dicionarios.html>

- erros comuns da língua portuguesa escrita.

Na Tabela 4.4 são apresentados alguns exemplos de erros com a sua identificação e descrição.

Tabela 4.4: Exemplos de erros do CoGrOO.

Erro	Descrição
xml:17	O adjetivo concorda com o substantivo a que se refere.
xml:67	Certos advérbios atraem o pronome para antes do verbo.
xml:75	Conjugação de um verbo irregular no futuro do subjuntivo.
xml:124	O adjetivo na função de predicativo concorda com o sujeito.
government	Regência verbal.
probs:paronyms	Possível confusão entre bebê e bebe.

Fonte: (Autor, 2021)

As manutenções e atualizações do CoGrOO ocorrem de forma colaborativa, por meio da sua comunidade de desenvolvedores e colaboradores¹⁶ que fazem correções e submetem sentenças que possam causar falha na análise do corretor, indicando o resultado esperado. Essas submissões são utilizadas para uma melhoria contínua do corretor, que possui uma lista diversos trabalhos nos quais é utilizado, pois além de se tratar de uma aplicação de uso livre, apresenta maturidade como ferramenta de suporte à pesquisas da área de PLN.

4.5 VINCULAÇÃO DE ENTIDADES

A coerência textual se estrutura em grande parte por meio dos mecanismos de referências, em especial por aqueles relacionados às entidades do texto que representam entidades do mundo real. Assim, detectar e extrair estas entidades tem papel fundamental no processo de avaliação aqui proposto. Esse processo de detecção envolve etapas distintas, sendo a primeira delas a identificação e extração destas entidades, por meio de um processo de detecção de menção à entidades.

Após a marcação das entidades, é preciso detectar ou agrupar as referências comuns. Porém, antes deste passo é comum ser executado um processo de desambiguação e vinculação de entidades para facilitar o processo de resolução de correferências. Para execução desta tarefa, pela ausência de uma ferramenta disponível, foi feita a criação de uma base de conhecimento baseado é um dump da Wikipedia para o português. O dump em formato JSON foi obtido a partir da base Wikimedia¹⁷, de onde foram obtidas uma lista de entidades e de referência equivalentes (alias).

Para cada alias de uma entidade, tem-se um valor associado à probabilidade daquela referência indicar determinada entidade. A Tabela 4.5 mostra duas representações de uma mesma entidade do mundo real, identificada pelo mesmo identificador na base de conhecimento (KB_ID) e os seus possíveis alias.

Uma vantagem do uso de dados de entidades e alias extraídos da Wikipedia, além da cobertura dos dados, é a possibilidade de relacionar entidades por meio das referências mais comuns, ainda que a representação das referências estejam com grafias distintas ou até mesmo, ortograficamente incorretas. A Tabela 4.6 traz um exemplo de duas referências distintas à entidade

¹⁶<http://comunidade.cogroo.org>

¹⁷<https://dumps.wikimedia.org>

Tabela 4.5: Exemplo de Entradas da Base de Conhecimento.

REPRESENTAÇÃO	DADOS	CANDIDATOS
Banco Central	Entidade: LOC	Banco Central (similaridade = 1,00)
	KB_ID: Q8061	Banco Central do Brasil (similaridade = 0,81)
BC	Entidade: LOC	BC (similaridade = 1,00)
	KB_ID: Q806186	
Banco Central do Brasil	Entidade: LOC	Banco Central do Brasil (similaridade = 1,00)
	KB_ID: Q806186	Bacen (similaridade = 0,81)

Fonte: (Autor, 2021)

“vírus Zika” em que uma delas contém uma grafia incorreta. Essas diferentes representações serão considerados quando da avaliação dos elos coesivos, pois o modelo de vinculação de entidades irá considerar se tratar da mesma entidade.

Tabela 4.6: Entidade com duas representações similares.

REPRESENTAÇÃO	DADOS	CANDIDATOS
Zica Vírus	Tipo de Entidade: UNK	Zica Vírus (similaridade = 1,00)
	KB_ID: Q202864	Vírus da Zica (similaridade = 0,96)
Zika Vírus	Tipo de Entidade: UNK	Zika Vírus (similaridade = 1,00)
	KB_ID: Q202864	Zika (similaridade = 0,91)

Fonte: (Autor, 2021)

No processo de leitura, o leitor não deve ser considerado como um agente passivo, mas sim um ente que cria um caminho próprio na leitura, fazendo previsões sobre sentidos e temas, integrando os elementos textuais com elementos externos, que por meio de um processo metacognitivo cria os elos entre o texto e o mundo que o autor possa não ter deixado explícito (Lima, 2017). Assim, ao ler um texto que contenha incorreções mínimas, o leitor, a depender da sua capacidade, consegue reestabelecer ou fortalecer elos coesivos.

4.6 RESOLUÇÃO DE CORREFERÊNCIA

De acordo com as discussões apresentadas no Capítulo 2, as relações de encadeamento de referências tem grande importância no estabelecimento da coesão. Assim, avaliar a qualidade coesiva de um texto, exige a correta identificação de todos os termos que fazem referência a uma mesma entidade dentro do texto. Esses elementos são ditos correferentes e formam um conjunto de termos referenciado como cadeia de correferência.

Há diversos métodos propostos para a tarefa de resolução de correferência, dentre as quais podemos destacar:

- Mention-pair: utilizam modelos que são treinados a partir de pares de menções;
- Entity-mention: trabalha com pares, mas identifica entidades novas ao discursos e aquelas que são anafóricas;
- Mention-ranking: cria um *ranking* entre os candidatos, apresentando semelhanças com o Mention-pair;
- Neural mention-ranking: não separa o processamento em etapas de extração das menções ou de entidades, para depois gerar um ranking e executar um processo de seleção. Ao invés disso, utiliza uma rede neural para treinar um modelo que realiza a identificação das cadeias.

A despeito das diferentes propostas existentes para cada uma das arquiteturas citadas anteriormente, há poucas propostas para o português. Com os melhores resultados reportados para o português brasileiro tendo sido obtidos com o uso da ferramenta CORP (Fonseca et al., 2018b). O modelo proposto pelo CORP se utiliza de um conjunto de regras e define pesos para essas regras. Os tipos de correferências detectadas pelo modelo incluem:

Tabela 4.7: Lista de regras do CORP.

Código	Descrição
Casamento de Padrões Exato	[o Brasil], [o Brasil]
Casamento Parcial pelo Núcleo	[o ministro da justiça], [o ministro]
Aposto Explicativo	[A ministra da justiça do país], [Elisabete Guigou], ...
Aposto Especificativo	[o arqueólogo português], [Francisco Alves]
Acrônimo	[Organização das Nações Unidas], [a ONU]
Predicado Nominativo	[A França] é [o único país que se recusa a aceitar] ...
Pronome Relativo	[Wilkinson Microwave Anisotropy Probe], [cujos] primeiros ...
Casamento Restrito pelo Núcleo1	[Universidade Federal de São Paulo] ... [a Universidade] ...
Casamento Restrito pelo Núcleo2	[O menino de pijama listrado], [o pijama listrado].
Casamento entre Nomes Próprios	[Califórnia],[a região sul da Califórnia].
Casamento Parcial entre Nomes Próprios	[o agrônomo da UFSC, Miguel Guerra] e [Guerra]
Hiponímia	... como [as abelhas] fabricam mel? [Os insetos] saem ...
Sinonímia	[A Terra] é [um astro do sistema solar]. [Esse planeta] ...

Fonte: Adaptado de (Fonseca et al., 2018b)

Para a identificação das cadeias de correferências, o CORP extrai uma lista de menções e após análise, devolve uma lista de cadeias devidamente agrupadas, por meio das suas regras que levam em conta os pesos atribuídos. Gerando, inclusive, cadeias contendo apenas uma única entidade, sem nenhuma menção - singleton.

Para ilustrar a execução do CORP sobre um dos exemplares da base de textos, nas Tabelas 4.8 e 4.9 são apresentadas respectivamente, um exemplar de texto e suas cadeias de correferência detectadas pelo CORP ¹⁸.

Tabela 4.8: Texto separado em sentenças.

ID	Sentença
0	Embora as pessoas acreditem que a mulher só obteve seu papel nos últimos séculos, há registros de mulheres que se destacaram na antiguidade clássica, como é o caso da Cleópatra e Hatshepsut (rainhas do Egito), mulheres espartanas e mulheres romanas, apesar dessas não terem uma vida política ativa.
1	Houve progressos e retrocessos ao longo da história nesse âmbito.
2	No entanto, no Brasil, o que se vê é que as mulheres foram tardiamente e gradativamente conquistando seus direitos.
3	A partir de 1827, a mulher brasileira passou a ter brandos direitos, como permissão de estudar, e em 1888 destaca-se a ‘pioneira’ em ‘questões diplomáticas’, Princesa Isabel, que assinou a Lei Áurea, a qual aboliu a escravidão.
4	Apesar de ela ter um papel fundamental na história do Brasil, demorou a outras mulheres se destacarem e, após muitas lutas, Bertha Luz, em 1922, cria a Federação Brasileira pelo Progresso Feminino, que foi o estopim para a conquista do voto feminino.
5	A partir daí, a mulher começou a progredir um pouco mais rapidamente, tendo em 1935 a primeira parlamentar eleita, Carlota de Queiroz.
6	Conforme a constituição, todos são iguais perante a lei não havendo mais distinção de sexo, no entanto ainda há em pleno século XXI preconceitos machistas vindos tanto de homem como de mulheres.
7	Mas isso não é problema para abalar a boa realidade vivida pelas mulheres, já que elas são maioria no país e no eleitorado uma vez que 70 milhões de mulheres votam sendo cinco milhões a mais que os homens, segundo a Agência Nacional.
8	O sexo frágil com certeza já se tornou forte, não só na política, mas nas universidades as mulheres estão ascendendo cada vez mais em relação aos homens, e nas últimas décadas as mulheres está se matriculando 22% a mais que os homens, segundo Instituto Nacional de Estudos e Pesquisas Educacionais (Inep).
9	No entanto nessa maré de rosas, as mulheres sofrem com a discriminação quanto ao salário, pois muitas exercem cargos maiores que de homens e ganham salários inferiores.
10	Mas paulatinamente essas guerreiras vão enfrentar e, com certeza, vencer mais essa batalha!

Nota: Quebras de parágrafo entre os pares de sentenças (2; 3), (5; 6) e (7; 8)

Fonte: (Autor, 2021)

¹⁸No Apêndice B, além da identificação das cadeias de correferência, traz-se uma análise completa de um texto com todas as anotações e identificações de entidades, das listas de foco explícito e implícito.

Tabela 4.9: Tabela de Cadeias de Correferências.

Número da Cadeia	Correferências	
	Sentença	Sintagma
Cadeia 1	0	as pessoas
	0	a mulher
	5	a mulher
	6	homem
	7	os homens
	8	os homens
	9	homens
Cadeia 2	0	mulheres
	2	as mulheres
	6	mulheres
	7	as mulheres
	7	mulheres
	8	as mulheres
	9	as mulheres
Cadeia 3	0	o caso de a Cleópatra e Hatshepsut
	1	a história
	4	a história
Cadeia 4	0	Hatshepsut
	0	rainhas de o Egito
Cadeia 5	2	o Brasil
	4	o Brasil
Cadeia 6	4	ela
	4	um papel fundamental
Cadeia 7	6	século XXI preconceitos machistas
	7	maioria
Cadeia 8	8	certeza
	10	certeza
Cadeia 9	8	Pesquisas Educacionais
	8	Inep

Fonte: (Autor, 2021)

Os resultados do CORP foram avaliados utilizando os corpus anotados Summ-it++ e Corref-PT, comparando-os com outras propostas, em especial com a de Garcia (2014), que apresenta resultados superiores para determinados tipos de correferência, como aquelas relacionadas a pronomes pessoais. Contudo, na maioria dos casos, para os dois corpus utilizados, a proposta de (Fonseca et al., 2018b) demonstrou maior abrangência, sendo esta a principal motivação sua escolha para a tarefa de resolução de correferência neste trabalho.

4.6.1 Cálculo da Coesão

Nesta seção é apresentado um exemplo de execução do método para valoração de um conjunto de pares de sentenças, bem como a execução para valoração de um par de parágrafos. Para exemplificar o funcionamento do método proposto, considere o processamento do trecho de texto a seguir:

Na Grécia antiga, berço de grande parte da cultura ocidental contemporânea, a relação entre homens e animais era voltada para prática agrícola. Como o tempo, está foi tornando-se mais afetiva e até proporcionar certos mimos luxuosos para estes animais. Hoje, esses têm sido questionados pela sociedade vendo os cuidados e tratamentos para estes, sendo que na maior parte da população crianças e adolescentes passam fome e sofrem pelo abandono. Quem quiser compreender - e modificar - essa realidade deverá analisar as causas e suas consequências.

Antes e tudo, todas as crianças e adolescentes têm o direito de ser amados e protegidos em todos os sentidos. De fato, a fome é a grande vilão da alta taxa mortalidade infantil e também pelo retardamento do desenvolvimento mental e físico, - demonstrando a indiferença do homem com o próprio homem -. Isso talvez explique a valorização de um animal de estimação e a degradação do sentimento com o próximo.

Para agravar a situação, os dados estatísticos revelam crescimento de crianças e adolescentes abandonadas nas ruas. Uma pesquisa realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE) relata que: "Em 2009 houve o aumento de 23% (vinte três por cento) de crianças e adolescentes que vivem nas ruas das principais metrópoles do Brasil", está triste notícia demonstra como estes são desprovidos de qualquer qualidade de vida e também ausência dos direitos humanos. Por conseguinte, retrata a total falta de - amor - com o seu semelhante.

O processamento necessário para avaliação de um texto se inicia com a segmentação do texto em suas sentenças. Seguido pela anotação das informações sintáticas das partes componentes do texto pelo parser do Spacy. Em paralelo é feita a anotação das informações semânticas pelo parser PALAVRAS. Como os dois parsers trabalham com anotações distintas é necessário alinhar os elementos e anotações de cada um dos tokens do texto, gerando um objeto - documento texto - que contenha tanto as informações do Spacy, quanto do PALAVRAS.

Uma vez que o processo de anotações das informações sintáticas e semânticas foi feito, executa-se os processos de menção de entidades e reconhecimento de entidades nomeadas. Faz uma avaliação de quais entidades podem ter sido detectadas em duplicidade pelos dois processos, eliminando as repetições. Uma vez concluído este processamento, passa-se para a avaliação dos pares de sentenças.

4.6.1.1 Avaliação dos Pares de Sentenças

As sentenças são analisadas aos pares, avaliando as relações entre as suas listas de foco explícito e implícito, e assim é calculada uma nota para a relação delas. A lista de foco explícito é construída utilizando informações de dicionário, informações sintáticas e as entidades reconhecidas. Já a lista de foco implícito é construída a partir das informações semânticas. O resultado do processamento inicial de cada uma das sentenças do primeiro parágrafo do texto é apresentado na Tabela 4.10.

Uma vez anotadas as informações de cada uma das sentenças, é feita a análise destas informações considerados os pares subsequentes de sentenças. Assim, serão analisados os pares (0,1), (1,2) e (2,3).

Durante a análise do par formado pelas sentenças 0 e 1, percebe-se que duas referências similares são detectadas: Grécia e Grécia Antiga. Para validar se essas referências estão vinculadas a uma mesma entidade, podemos utilizar a informações de vinculação de entidades, que nos apresenta a seguinte informação: {Grécia: {label: 385, kb_id: Q41, candidates: [Grécia Antiga, Grécia Clássica]}. O que indica que as referências podem ser consideradas equivalentes. Por conta disso, somente uma delas figura na lista de Foco Explícito.

Tabela 4.10: Anotação de Informações das Sentenças

Sentença 0	Na Grécia antiga, berço de grande parte da cultura ocidental contemporânea, a relação entre homens e animais era voltada para prática agrícola.
Entidades Nomeadas	{ Grécia }
Menções DBPedia	{ cultura ocidental, Grécia antiga }
Foco Explícito	{ homens, prática, Grécia antiga, parte, cultura ocidental, berço, relação, animais }
Foco Implícito	{ berço:[<furn>, <Lh>, <tool>] parte:[<part>, <amount>, <L>, <HH>, <sem-l>, <ac>] cultura:[<ac>, <activity>] relação:[<ac>, <sem-r>] homens:[<Hbio>] animais:[<Azo>] prática:[<per>, <sit>, <act-d>, <talk>] }
Sentença 1	Como o tempo, está foi tornando-se mais afetiva e até proporcionar certos mimos luxuosos para estes animais.
Entidades Nomeadas	{ }
Menções DBPedia	{ mimos }
Foco Explícito	{ mimos, tempo, animais }
Foco Implícito	{ tempo:[<am>, <temp>, <per>, <ac-cat>, <wea>] }
Sentença 2	Hoje, esses têm sido questionados pela sociedade vendo os cuidados e tratamentos para estes, sendo que na maior parte da população crianças e adolescentes passam fome e sofrem pelo abandono.
Entidades Nomeadas	{ }
Menções DBPedia	{ }
Foco Explícito	{ abandono, fome, população, crianças, parte, sociedade, cuidados, estes, adolescentes, tratamentos, sido, esses }
Foco Implícito	{ sociedade:[<HH>] cuidados:[<am>, <act-d>, <ac>] tratamentos:[<therapy>, <f-c>, <activity>] parte:[<part>, <amount>, <L>, <HH>, <sem-l>, <ac>] população:[<HH>] crianças:[<H>] adolescentes:[<Hbio>] fome:[<am>] abandono:[<act>, <state>] }
Sentença 3	Quem quiser compreender - e modificar - essa realidade deverá analisar as causas e suas conseqüências.
Entidades Nomeadas	{ }
Menções DBPedia	{ }
Foco Explícito	{ realidade, conseqüências, causas, quem }
Foco Implícito	{ realidade:[<vH>] }

Fonte: (Autor, 2021)

Considerando a primeira e segunda sentenças do texto, ao calcular-se a interseção das listas de foco explícito das duas sentenças, tem-se o seguinte conjunto: { animais }. Já ao calcular a interseção das listas de foco implícito das duas sentenças, tem-se o seguinte conjunto: { <per> }. Considerando a tabela de valoração das relações, apresentada no Capítulo 2, tem-se que a nota da coesão local será 1,0. O que corresponde a uma transição de elaboração entre as duas sentenças. O processo de análise apresentado deve ser reproduzido para todos os pares de sentenças subsequentes, considerando o valor para cada tipo de transição entre os pares. Uma vez executada a análise para todos os pares que foram anteriormente anotados, tem-se o resultado apresentado na Tabela 4.11.

Tabela 4.11: Valor da Relação dos Pares de Sentenças

Par de Sentenças:	(0, 1)
Interseção Foco Explícito:	{ animais }
Interseção Foco Implícito:	{ per }
Valor da Relação:	1,00
Par de Sentenças:	(1, 2)
Interseção Foco Explícito:	{ }
Interseção Foco Implícito:	{ am }
Valor da Relação:	0,50
Par de Sentenças:	(2, 3)
Interseção Foco Explícito:	{ }
Interseção Foco Implícito:	{ }
Valor da Relação:	0,00

Fonte: (Autor, 2021)

De posse dos resultados para os pares de sentenças subsequentes, para calcular o valor total da coesão local basta somar os valores obtidos e dividir pelo número de pares formados. Assim, para o trecho de texto analisado o valor da coesão local será $(1,00 + 0,50 + 0,00) / 3 = 0,50$. Este valor será utilizado para compor o índice de coesão do texto como um todo, quando somado ao valor da coesão global, obtida a partir da análise dos pares de parágrafos.

4.6.1.2 Avaliação dos Pares de Parágrafos

A avaliação dos parágrafos é feita, da mesma maneira que nos pares de sentenças, considerando as listas de foco explícito e implícito e avaliando as suas interseções. Entretanto, as listas são construídas considerando as listas já formadas pelas sentenças componentes dos parágrafos e os pares a serem avaliados são feitos de forma a combinar todos os parágrafos entre si. Obtendo-se assim todos os pares possíveis, sem repetição. As listas de foco explícito e implícito dos parágrafos são construídas a partir da união das interseções das respectivas listas dos pares de sentenças.

Assim, ao avaliar os parágrafos, serão utilizadas as listas de foco implícito e explícito das sentenças que os compõem. Contudo, as listas não serão utilizadas de forma integral e a combinação de pares de sentenças consideradas, não serão apenas as combinações sub-sequentes. Para ilustrar melhor, considere-se um parágrafo P qualquer, composto por 3 sentenças, $S1$, $S2$ e $S3$. A lista de foco explícito deste parágrafo será formada pela união das interseções das listas de foco explícito dos pares que as suas sentenças podem formar. De maneira que:

- Sentenças: $S1$, $S2$ e $S3$
- Focos Explícitos das sentenças: FE_{S1} , FE_{S2} e FE_{S3}
- Pares possíveis: (FE_{S1}, FE_{S2}) , (FE_{S1}, FE_{S3}) , (FE_{S2}, FE_{S3})
- Foco Explícito do Parágrafo: $(FE_{S1} \cap FE_{S2}) \cup (FE_{S1} \cap FE_{S3}) \cup (FE_{S2} \cap FE_{S3})$

Para obtenção da lista de foco implícito, o mesmo raciocínio da listas de foco explícito é aplicado. Assim, a avaliação dos parágrafos pode ser executada. Um exemplo de avaliação de 03 parágrafos distintos é apresentado na Tabela <>.

Tabela 4.12: Avaliação dos Pares de Parágrafos

Parágrafo 0	Na Grécia antiga, berço de grande parte da cultura ocidental contemporânea, a relação entre homens e animais era voltada para prática agrícola. Como o tempo, está foi tornando-se mais afetiva e até proporcionar certos mimos luxuosos para estes animais. Hoje, esses têm sido questionados pela sociedade vendo os cuidados e tratamentos para estes, sendo que na maior parte da população crianças e adolescentes passam fome e sofrem pelo abandono. Quem quiser compreender - e modificar - essa realidade deverá analisar as causas e suas consequências .
Foco Explícito	{ parte, animais }
Foco Implícito	{ parte: [<part>, <amount>, <L>, <HH>, <sem-l>, <ac>], cultura: [<ac>, <activity>], relação: [<ac>, <sem-r>], homens: [<Hbio>], prática: [<per>, <sit>, <act-d>, <talk>], tempo: [<am>, <temp>, <per>, <ac-cat>, <wea>], cuidados: [<am>, <act-d>, <ac>], tratamentos: [<therapy>, <f-c>, <activity>], adolescentes: [<Hbio>], fome: [<am>] }
Parágrafo 1	Antes e tudo, todas as crianças e adolescentes têm o direito de ser amados e protegidos em todos os sentidos. De fato, a fome é a grande vilão da alta taxa mortalidade infantil e também pelo retardamento do desenvolvimento mental e físico, - demonstrando a indiferença do homem com o próprio homem -. Isso talvez explique a valorização de um animal de estimação e a degradação do sentimento com o próximo.
Foco Explícito	{ }
Foco Implícito	{ retardamento: [<act>], desenvolvimento: [<process>, <act>], valorização: [<act>], degradação: [<act>, <process>] }
Parágrafo 2	Para agravar a situação, os dados estatísticos revelam crescimento de crianças e adolescentes abandonadas nas ruas. Uma pesquisa realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE) relata que: "Em 2009 houve o aumento de 23% (vinte três por cento) de crianças e adolescentes que vivem nas ruas das principais metrópoles do Brasil", está triste notícia demonstra como estes são desprovidos de qualquer qualidade de vida e também ausência dos direitos humanos. Por conseguinte, retrata a total falta de \- amor - com o seu semelhante.
Foco Explícito:	{ adolescentes, crianças, ruas }
Foco Implícito:	{ situação: [<sit>], falta: [<f-c>, <sit>, <event>] }

Fpnte: (Autor, 2021)

Tabela 4.13: Resultados da Avaliação dos Pares de Parágrafos

Valor da Relação dos Pares de Parágrafos	
Par de Parágrafos	(0, 1)
Interseção Foco Explícito	{ }
Interseção Foco Implícito	{ }
Valor da Relação	0,00
Par de Parágrafos	(0, 2)
Interseção Foco Explícito	{ }
Interseção Foco Implícito	{ f-c, sit }
Valor da Relação	0,50
Par de Sentenças	(1, 2)
Interseção Foco Explícito	{ }
Interseção Foco Implícito	{ }
Valor da Relação	0,00

Fonte: (Autor, 2021)

Uma vez executada a análise para todos os pares de parágrafos, os mesmos são avaliados aos pares. Diferentemente das sentenças cujos pares são formados de maneira sequenciada, para os parágrafos são considerados todos os pares possíveis, sem repetição. Assim, ao analisarmos os 03 parágrafos apresentados anteriormente, considerando as suas listas de foco explícito e implícito, bem como a tabela de valoração das transições apresentadas no Capítulo 2, tem-se o resultado apresentado na Tabela 4.13.

De posse dos resultados para os pares de parágrafos, para calcular o valor total da coesão global basta somar os valores obtidos e dividir pelo número de pares formados. Assim, para o trecho de texto analisado o valor da coesão global será $(0,00 + 0,50 + 0,00)/3 = 0,17$. Este valor será utilizado para compor o índice de coesão geral do texto, quando somado ao valor da coesão local, obtida a partir da análise dos pares de subsequentes de sentenças.

De posse dos valores da coesão global (C_g) e local (C_l), o Índice de Coesão do texto será obtido, conforme apresentado no Capítulo 2, da seguinte forma $(C_g + C_l)/2$. Contudo, outros fatores podem contribuir para a coesão do texto. Como o Índice de Forma (IF) e o Índice de Ambiguidades (IAC), que podem ser obtidos por contagens diretas de elementos identificados pelo corretor ortográfico CoGrOO e calculados conforme suas respectivas equações apresentadas no Capítulo 2. De posse dos 03 índices citados, é possível inferir um valor de nota relativa à coesão textual, utilizando-se, por exemplo, um método de regressão supervisionada e a avaliação da qualidade da predição pode ser feita por meio do Kappa Médio Quadrático anteriormente apresentado. O que foi feito durante os experimentos aqui executados.

4.6.2 Comparação com Outros Trabalhos

Durante a execução dos experimentos, também se buscou repetir a forma de avaliação utilizada em Nobre (2011). Todavia, não foi utilizado o mesmo método de inferência de valores, que no caso do estudo citado foram usados métodos de lógica fuzzy. A utilização destes métodos foi substituída pelo uso de modelos de regressão, pois estes permitem aplicação direta para outras bases de texto, sem exigir esforços de ajustes das funções de inferência ou calibrações. Além

disso, os métodos regressivos apresentaram bons resultados nos trabalhos com os quais foram feitas comparações dos resultados.

Para avaliar os resultados obtidos nos experimentos conduzidos, foram feitas comparações utilizando as métricas QWK e RMSE, com os trabalhos de Amorim e Veloso (2017) e Fonseca et al. (2018) com relação aos valores da Competência 4. Ainda que a base de textos utilizada por Fonseca et al. seja composta por textos de redações do ENEM, consideramos que seja válida a comparação. No caso dos textos de Amorim e Veloso, a fonte dos dados e o tipo de texto é bastante similar, além dos seus resultados servirem de base para outros estudos.

Para as comparações, também se considerou os resultados apresentados por Marinho et al. (2021), ainda que este não tenha necessariamente apresentado uma proposta de avaliação, mas sim repetido os métodos de Fonseca et al. e Amorim e Veloso em uma nova base de textos dissertativos. Esta base possui uma quantidade de textos - assim como a nossa base - muito inferior à base de Fonseca et al., mas é consideravelmente maior que a base utilizada por Amorim e Veloso. Desta maneira, acreditamos que a comparação dos valores obtidos nesta base com os valores obtidos pelo método aqui apresentado seja o mais adequado, por conta similaridade dos textos e por conta da quantidade de redações.

4.7 CONCLUSÃO

Neste capítulo apresentou-se a arquitetura da solução de avaliação e valoração da coesão proposta, bem como as principais ferramentas utilizadas para as diferentes tarefas necessárias ao processo de avaliação e valoração de coesão textual. Ao final do capítulo traz-se as comparações feitas com outros trabalhos com os quais os valores obtidos pelo protótipo desenvolvido foram comparados durante a execução dos experimentos apresentados no próximo capítulo.

5 EXPERIMENTOS E RESULTADOS

Neste capítulo será apresentado o conjunto de dados utilizado nas análises e serão descritos e discutidos os experimentos realizados, por meio do protótipo desenvolvido, como forma de comprovar o potencial do método proposto.

5.1 BASE DE TEXTOS

Os textos que compõem a base de redações foram extraídos do portal Brasil Escola¹, onde é mantido um banco de redações enviadas por alunos do ensino médio que são avaliadas por corretores especializados em redações dissertativas. Estes textos estão disponíveis de forma pública e são enviados, em sua maioria, por alunos inscritos em programas voltados para melhoria da escrita. Após o processo de *scrapping* efetuado, foram executados os devidos tratamentos para extração dos textos e valores a partir de documentos em formato HTML. Alguns textos tiveram que ser excluídos do conjunto inicial, pois apresentavam problemas que aparentemente estavam relacionados à falhas no processo de envio ou preparação das redações. Assim, foram desconsiderados textos que receberam nota total zero por fuga ao tema. Também foram desconsiderados os textos que não tiveram a sua correção executada e receberam nota 0 em todas as suas competências, por não serem de utilidade para o processo de comparação.

No processo de extração dos textos também foram coletados os temas a eles relacionados, porém esta informação não será considerada, pois avaliação de adequação ao tema, que é considerado como uma competência na correção de redações, mesmo estando presente em métodos de AAR, foge ao escopo deste estudo. Dessa maneira, na valoração dos textos foram consideradas 5 competências esperadas do escritor, onde cada uma delas aborda aspectos distintos do processo de escrita. Segundo o critério utilizado pela fonte de obtenção das redações, que segue as mesmas especificações do MEC relativas à redação do ENEM, as competências são valoradas em uma escala de 0 a 200 pontos e para que um texto receba valoração máxima, ele deve atender, em cada uma das competências, os seguintes requisitos:

- Competência 1 (Ortografia): Demonstrar domínio da norma culta da língua escrita;
- Competência 2 (Adequação ao tema): Compreender a proposta da redação e aplicar conceitos de várias áreas de conhecimento para desenvolver o tema, dentro dos limites estruturais do texto;
- Competência 3 (Coerência): Selecionar, relacionar, organizar e interpretar informações, fatos, opiniões e argumentos em defesa de um ponto de vista;
- Competência 4 (Coesão): Demonstrar conhecimento dos mecanismos linguísticos necessários para a construção da argumentação;
- Competência 5 (Ética): Elaborar a proposta de solução para o problema abordado, mostrando respeito aos valores humanos e considerando a diversidade sociocultural.

Faz-se necessário destacar que textos com nota zero na Competência 4, mas com nota final e nota definida para as demais competências foram mantidos, sendo inclusive considerados

¹<https://brasilecola.uol.com.br/redacao>

úteis ao processo de avaliação, ainda que a quantidade de instâncias desse tipo na base seja pequena. Após todo o processo de tratamento dos dados, obteve-se um total de 2.600 textos, relacionados a diversos temas, com as suas respectivas correções, notas nas competências técnicas de escrita e uma nota total, resultado da somatória das notas das competências. A Tabela 5.1 traz informações sobre os textos da base, seus tokens, sentenças, parágrafos.

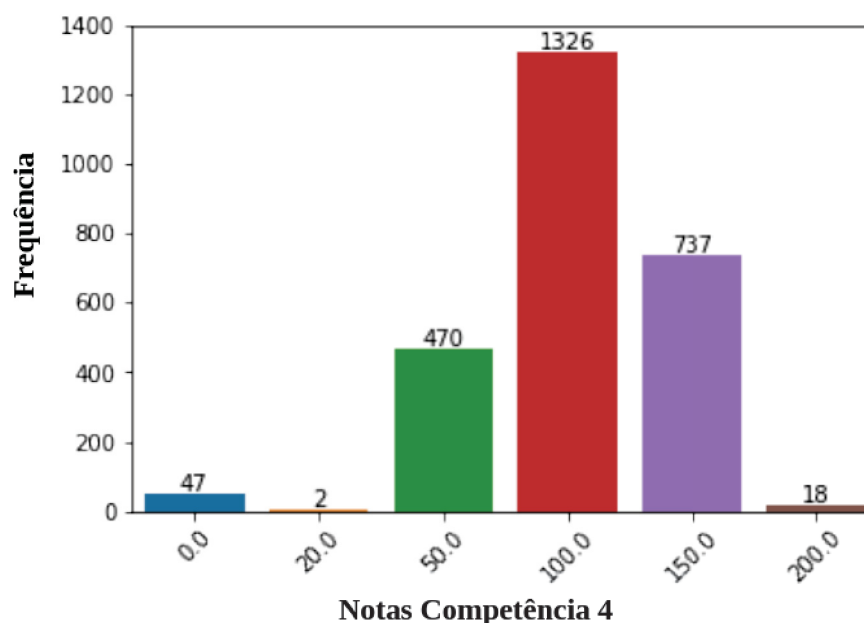
Tabela 5.1: Dados do conjunto de redações.

Relação	Taxa	Desvio-padrão
parágrafos/redação	4,42	0,92
sentenças/parágrafo	2,63	1,37
sentenças/redação	11,61	3,92
tokens/sentença	29,97	16,79
tokens/parágrafo	78,74	29,66
tokens/redação	348,11	56,33

Fonte: (Autor, 2021)

Neste estudo nos interessa os valores da Competência 4 que estão diretamente relacionados à coesão textual. Somente esta competência será utilizada para fins de comparação dos nossos resultados. Na Figura 5.1 exibe-se a distribuição das notas da Competência 4 em toda a base de textos. Sobre a distribuição das notas da Competência 4 foi aplicado um teste de normalidade do tipo K^2 de D'Agostino, que mostrou que os dados não seguem uma distribuição normal, tendo-se obtidos valores de probabilidade de 82.781 para um valor de $p = 0.000$.

Figura 5.1: Distribuição das notas da Competência 4.



Fonte: [Autor, 2021]

Ao optar por analisar a coesão de forma isolada busca-se garantir que atributos ou dados relacionados à outras competências não impactem na avaliação realizada, o que, caso aconteça,

torna difícil determinar quais os atributos que causaram esse impacto e qual o seu nível. Fato é que em trabalhos anteriores, como por exemplo (Fonseca et al., 2018a), os autores reportaram que modelos bons em valorar uma competência, tendem a ser ruins nas demais, efeito que se buscou evitar tanto com relação às outras competências, quanto com relação algumas das partes que compõem os processos de identificação de entidades e de valorações presentes no modelo aqui desenvolvido. A análise do impacto de diferentes partes do método foi feita por meio dos experimentos apresentados na próxima seção.

5.2 EXPERIMENTOS

Nesta seção são apresentados os experimentos executados sobre a base de redações apresentada na seção 5.1. Grande parte do trabalho que precedeu os experimentos teve como foco a validação do método e das adaptações feitas, sendo necessário escolher algumas amostras de textos e executar uma revisão dos resultados entregues pelo método. O Apêndice B traz um exemplo de um texto com as anotações feitas sobre o mesmo e as informações utilizadas para avaliar o desempenho e a validade do método. Estas avaliações serviram como base para as correções no método aqui apresentado e na sua implementação.

5.2.1 Experimento 1: Valoração com Entidades e Referências

Neste experimento, considerando as etapas do modelo proposto, foram executadas as seguintes etapas:

1. Anotação de dados
 - Anotação das informações sintáticas e gramaticais pelo parser do Spacy
 - Reconhecimento de Entidades Nomeadas pelo parser do Spacy
 - Anotação das informações semânticas pelo parser PALAVRAS
 - Detecção das menções de entidades usando o DBPedia Spotlight
2. Agrupamento e alinhamento das anotações em relação ao texto original gerando um Objeto Documento
3. Agregação de resultados e exclusão de informações redundantes em um Objeto Análise Documento
4. Análise baseada na Teoria do Centragem e na Teoria do Foco
 - Busca de sinônimos das entidades, menções e referências na base de sinônimos do TeP2
 - Aplicação das regras de uso das informações obtidas para decidir informações a serem mantidas e excluídas
 - Análise da coesão local pelo método de Nobre (2011)
 - Análise da coesão global pelo método de Nobre (2011)
5. Cálculo dos valores da coesão
 - Cálculo dos valores da coesão local
 - Cálculo dos valores da coesão global

- Cálculo do valor único para a coesão

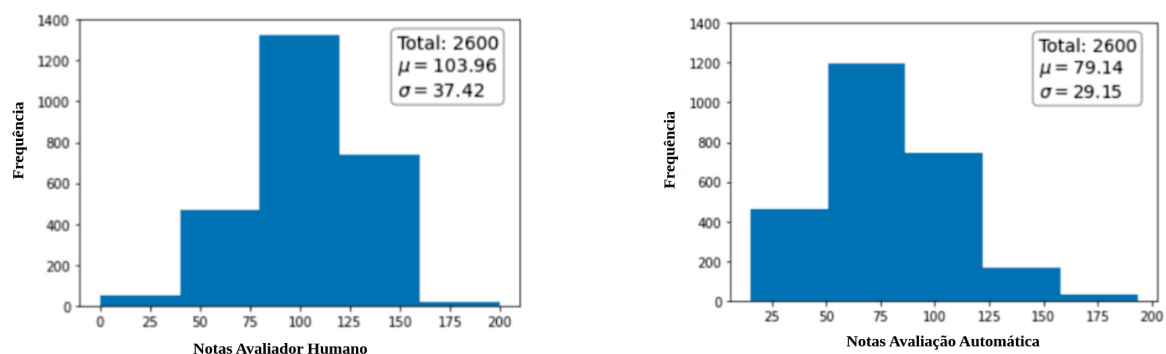
6. Comparação e cálculos com as notas

- Cálculo da diferença absoluta entre o valor da nota do avaliador humano e o valor calculado pelo método
- Cálculo da média dos valores das notas obtidos pelo método
- Cálculo dos valores de QWK e RMSE

7. Comparação dos valores do QWK e RMSE com trabalhos correlatos

Para estabelecer uma base de comparação foi realizado um experimento considerando apenas os valores obtidos por meio da Teoria do Foco e da Centragem, sem o uso dos métodos de resolução de correferência e vinculação de entidades. Esta execução nos permite ter uma noção do quão bem o método consegue avaliar um texto, utilizando apenas as entidades, suas referências diretas e as ligações semânticas entre elas. Na Figura 5.2 é apresentada a distribuição dos valores de coesão calculados pelo método em comparação com as notas do avaliador humano.

Figura 5.2: Comparação dos valores atribuídos pelo avaliador humano (esquerda) com os valores de coesão calculada (direita).

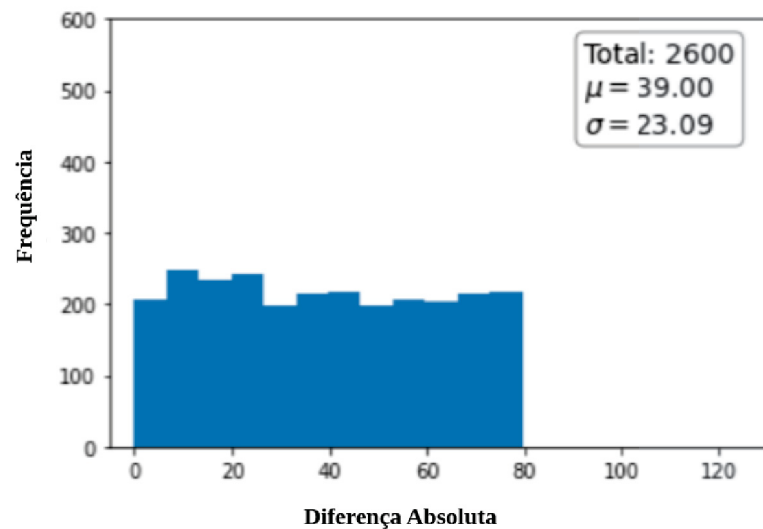


Fonte: (Autor, 2021)

Após o cálculo de um valor único para a coesão, englobando o valor da coesão local e global, conforme definido no Capítulo 2, é calculada, para cada texto da base, a diferença absoluta entre o valor da nota dada pelo avaliador e o valor calculado pelo método. Desta maneira, busca-se avaliar os valores médios e a distribuição do erro absoluto. Ao analisarmos a distribuição do valor absoluto da diferença, a média entre a nota do avaliador humano e a avaliação automática foi de 39 pontos, com desvio-padrão de 23,09. Enquanto o desvio-padrão apresentado pelo avaliador humano foi de 103,95. A diferença máxima da diferença absoluta foi de 79,49 pontos. Na Figura 5.3 é apresentada a distribuição das diferenças absolutas.

Chama a atenção do fato da avaliação automática, diferentemente do avaliador, não ter atribuído nota mínima (0.00), nem nota máxima (200.00) a nenhuma redação. Possivelmente, isso se dê por conta do método não conseguir capturar determinados aspectos da coesão em seu processo, uma vez que estamos utilizando apenas os valores obtidos por meio da Teoria da Centragem e do Foco.

Figura 5.3: Distribuição das diferenças absolutas das notas.



Fonte: (Autor, 2021)

5.2.2 Experimento 2: Valoração com Uso de Correferência

Neste experimento, considerando as etapas do modelo proposto, foram executadas as seguintes etapas:

1. Anotação de dados
 - Anotação das informações sintáticas e gramaticais pelo parser do Spacy
 - Reconhecimento de Entidades Nomeadas pelo parser do Spacy
 - Anotação das informações semânticas pelo parser PALAVRAS
 - Detecção das menções de entidades usando o DBPedia Spotlight
2. Agrupamento e alinhamento das anotações em relação ao texto original gerando um Objeto Documento
3. Processamento de Entidades
 - Execução do processo de vinculação de entidades
 - Execução do processo de resolução de correferências utilizando a ferramenta CORP.
4. Agregação de resultados e exclusão de informações redundantes em um Objeto Análise Documento
5. Análise baseada na Teoria do Centragem e na Teoria do Foco
 - Busca de sinônimos das entidades, menções e referências na base de sinônimos do TeP2
 - Aplicação das regras de uso das informações obtidas para decidir informações a serem mantidas e excluídas
 - Análise da coesão local pelo método de Nobre (2011)

- Análise da coesão global pelo método de Nobre (2011)

6. Cálculo dos valores da coesão

- Cálculo dos valores da coesão local
- Cálculo dos valores da coesão global
- Cálculo do valor único para a coesão

7. Comparação e cálculos com as notas

- Cálculo da diferença absoluta entre o valor da nota do avaliador humano e o valor calculado pelo método
- Cálculo da média dos valores das notas obtidos pelo método
- Cálculo dos valores de QWK e RMSE

8. Comparação dos valores do QWK e RMSE com trabalhos correlatos

Aqui as etapas do Experimento 1 foram novamente executadas, agora considerando a influência das correferências e da vinculação de entidades na construção das Listas de Foco Explícito e Implícito. Destacando-se que as correferências e a vinculação de entidades, são elementos que não constavam na proposta original do método de (Nobre, 2011). Na Tabela 5.2 é apresentada uma comparação entre os resultados obtidos nos experimentos 1a e 1b, com os valores das notas do avaliador humano.

Tabela 5.2: Valores de média, desvio-padrão e distribuição em quartis dos valores das notas dada pelo avaliador humano e das notas calculadas pelo método nos Experimentos 1 e 2.

	Nota Avaliador	Nota Experimento 1	Nota Experimento 2
Média	103,95	79,13	87,83
Desvio-padrão	37,42	29,15	30,62
Valor Mínimo	0,00	15,00	21,21
25%	100,00	58,93	65,62
50%	100,00	77,78	86,71
75%	150,00	95,83	106,67
Valor Máximo	200,00	193,75	200,00

Fonte: (Autor, 2021)

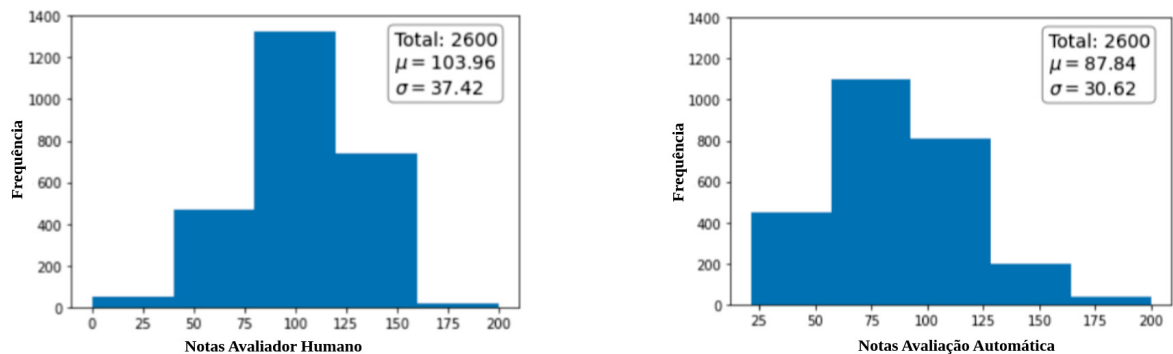
Com a inclusão das cadeias de correferências, o método foi capaz de dar notas no limite máximo de 200 pontos, sem contudo conseguir calcular notas de coesão iguais a zero. Neste caso, é provável que o problema não esteja relacionado à aspectos que o método não consiga capturar, uma vez que a divergência ocorre apenas quando o valor da nota é zero. Um provável motivo pode ser obtido na leitura da análise da Teoria da Centragem feita em (Taboada e Wieseemann, 2010), onde é possível encontrar uma citação à uma quinta possível transição dos centros de atenção, não explorado por Nobre (2011). Nesta transição - referenciada como transição nula - não é possível determinar em uma sentença o centro do foco de atenção com relação à sentença anterior à ela. Não sendo possível determinar como ocorre a relação de coesão entre essas sentenças.

No método de valoração de Nobre (2011) os pares que recebem valor zero são aqueles cuja transição encerra uma mudança de assunto porém, nas sentenças e parágrafos onde isso

acontece ainda é possível determinar centros de atenção. Este pode ser o motivo do método não dar nota zero para alguns textos. Contudo, também não se pode descartar a possibilidade que o valor zero do avaliador seja concebido por considerações subjetivas feitas por ele.

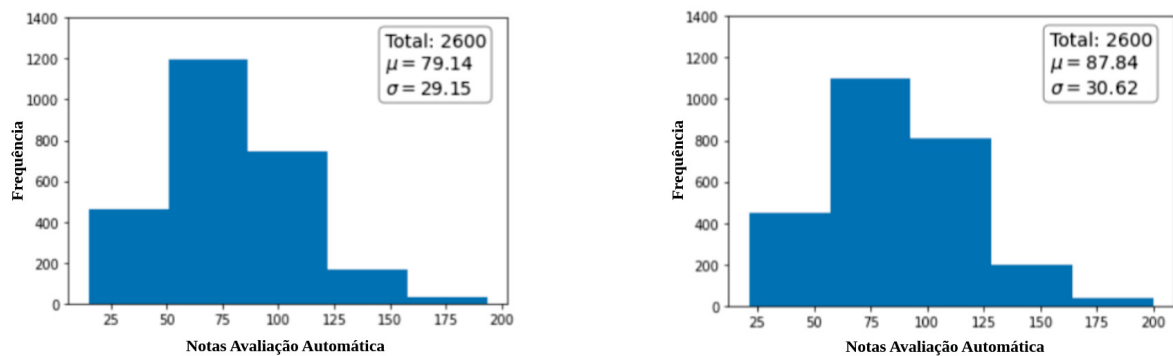
Com a inclusão das cadeias de correferências, nota-se uma mudança na distribuição das diferenças absolutas, conforme apresentado na Figura 5.4 onde tem-se uma comparação entre o último experimento e a distribuição das notas do avaliador humano e na Figura 5.5 uma comparação entre os resultados dos dois experimentos.

Figura 5.4: Comparação dos valores atribuídos pelo avaliador humano (esquerda) com os valores de coesão calculada pelo método proposto no Experimento 2 (direita).



Fonte: (Autor, 2021)

Figura 5.5: Comparação dos valores de coesão calculados no Experimento 1 (esquerda) e Experimento 2 (direita) pelo método proposto.



Fonte: (Autor, 2021)

A partir da análise dos resultados do experimento pode-se perceber uma redução na quantidade de erros com as maiores diferenças absolutas. Contudo, erros de magnitude superiores a 100 pontos surgem onde antes se tinha uma diferença absoluta máxima de 79,49 pontos, aumentando a dispersão das notas, porém ainda abaixo dos valores de dispersão de referência.

Para finalizar o Experimento 2, calculou-se os resultados das métricas do QWK e do RMSE, comparando-se os resultados dos experimentos com os resultados reportados por Amorim e Veloso em (Amorim e Veloso, 2017), Fonseca et al. em (Fonseca et al., 2018a) e Marinho

et al. (Marinho et al., 2021), para as notas da Competência 4. Para Marinho et al., apesar dos autores apresentarem duas abordagens distintas, é apresentado apenas o melhor resultado. A comparação dos valores do RMSE e do QWK pode ser vista na Tabela 5.3.

Tabela 5.3: Valores de QWK, quanto maior, melhor. Valores de RMSE, quanto menor, melhor. (1) Resultados utilizando o método de Amorim e Veloso. (2) Resultados utilizando o método de Fonseca et al.

Modelo	QWK	RMSE
Experimento 1	0,281	45,58
Experimento 2	0,312	42,39
Fonseca et al.	0,619	28,34
Marinho et al.	0,450	40,48
Amorim & Veloso	0,270	39,47

Fonte: (Autor, 2021)

Os resultados obtidos nos Experimentos 1 e 2 foram comparados segundo um teste t de Student bicaudal, considerando uma limiar de significância $p = 0.05$, tendo resultado em um valor de $p = 0.00$. Este resultado indica que ao nível de significância escolhido, ao considerar as cadeias de correferência e os meios de identificação de entidades do Experimento 2, agrega uma melhora real à avaliação da coesão do texto.

As comparações dos valores nos experimentos com os valores de outros trabalhos, mostram que os resultados obtidos no experimento, no melhor dos casos, se compara aos resultados do RMSE de Marinho et al., aproximando-se de Amorim e Veloso. Contudo, quando comparados considerando o QWK, os experimentos, com ou sem cadeias de correferências e vinculações, superam Amorim e Veloso e se aproximam do resultado de Marinho et al. Todavia, os valores obtidos nas duas métricas se mostram muito distantes do resultados obtidos por Fonseca et al., que é o melhor dentre todos os resultados considerados.

5.2.3 Experimento 3: Valoração com Índice de Forma e de Imprecisão

Neste experimento, considerando as etapas do modelo proposto, foram executadas as seguintes etapas:

1. Anotação de dados
 - Anotação das informações sintáticas e gramaticais pelo parser do Spacy
 - Reconhecimento de Entidades Nomeadas pelo parser do Spacy
 - Anotação das informações semânticas pelo parser PALAVRAS
 - Detecção das menções de entidades usando o DBPedia Spotlight
2. Agrupamento e alinhamento das anotações em relação ao texto original gerando um Objeto Documento
3. Processamento de Entidades
 - Execução do processo de vinculação de entidades
 - Execução do processo de resolução de correferências utilizando a ferramenta CORP.
4. Agregação de resultados e exclusão de informações redundantes em um Objeto Análise Documento

5. Análise baseada na Teoria do Centragem e na Teoria do Foco

- Busca de sinônimos das entidades, menções e referências na base de sinônimos do TeP2
- Busca dos erros gramaticais e ortográficos por meio do corretor CoGrOO
- Aplicação das regras de uso das informações obtidas para decidir informações a serem mantidas e excluídas²
- Análise da coesão local pelo método de Nobre (2011)
- Análise da coesão global pelo método de Nobre (2011)

6. Cálculo dos valores da coesão

- Cálculo dos valores da coesão local
- Cálculo dos valores da coesão global
- Cálculo do valor único para a coesão

7. Cálculo do Índice de Imprecisão

8. Cálculo do Índice de Forma

9. Separação dos dados em teste e treino

10. Treinamento de Modelo Regressivo utilizando a biblioteca scikit-learn

11. Teste do modelo regressivo

12. Comparação e cálculos com as notas

- Cálculo da diferença absoluta entre o valor da nota do avaliador humano e o valor calculado pelo método
- Cálculo da média dos valores das notas obtidos pelo método
- Cálculo dos valores de QWK e RMSE

13. Comparação dos valores do QWK e RMSE com trabalhos correlatos

No método de Nobre, diferente da abordagem feita nos Experimentos 1 e 2, são utilizados 3 índices para compor a nota de coesão. Além do Índice de Coesão (IC) são utilizados o Índice de Imprecisão (IAC) e o Índice de Forma (IF). Buscando reproduzir um método semelhantes ao método original, os índices foram calculados de acordo com a valoração destes mostrada no Capítulo 2. Cada um dos valores calculados foram fornecidos para diferentes modelos de regressão linear para avaliar qual se adequaria melhor. Os dados foram divididos em treinamento e teste, obedecendo a proporção 75-25 e os modelos foram testados utilizando um método de validação cruzada com 5 partições nos dados de treinamento. A escolha dos melhores modelos baseou-se no valor negativo da raiz quadrada do erro médio quadrático e no desvio-padrão dos valores obtidos nas partições. O resultado pode ser visto na Tabela 5.4 que traz os valores para cada um dos modelos.

²No Apêndice A são apresentadas algumas da regras de separação dos resultados fornecidos pelo CoGroo. As demais regras dizem respeito a eliminação de duplicidades na detecção de entidades, onde detecções repetidas ou partes de entidades são excluídas do processo.

Tabela 5.4: Resultado da avaliação de diferentes modelos de regressão, ordenados do menor para o maior.

Modelo	RMSE Negativo	Desvio-padrão
CART	-41,22	0,58
ExtraTreesRegressor	-35,09	1,06
KNN	-34,72	1,42
Regressão Linear	-34,06	1,67
LASSO	-34,09	1,65
Elastic Net	-34,47	1,71
GBM	-35,09	1,78

Fonte: (Autor, 2021)

Com base na avaliação da validação cruzada, os modelos baseados em CART, KNN e ExtraTreesRegressor foram executados sobre os dados de teste e foram obtidos os valores de RMSE e QWK apresentados na Tabela 5.5, bem como os valores obtidos no Experimento 1 e dos trabalhos utilizados para comparação.

Tabela 5.5: Valores de QWK, quanto maior, melhor. Valores de RMSE, quanto menor, melhor. (1) Resultados utilizando o método de Amorim Veloso. (2) Resultados utilizando o método de Fonseca et al.

Modelo	QWK	RMSE
ExtraTreesRegressor	0,568	32,83
CART	0,515	37,45
KNN (k=5)	0,396	36,39
Experimento 1a	0,281	45,58
Experimento 1b	0,312	42,39
Fonseca et al.	0,619	28,34
Marinho et al.	0,450	40,48
Amorim & Veloso	0,270	39,47

Fonte: (Autor, 2021)

Neste experimento também avaliou-se o quanto a inclusão das correferências e menção de entidades contribui para a avaliação utilizando modelo regressivo. Os valores obtidos são apresentados na Tabela 5.6, corroborando a avaliação feita a partir da comparação do Experimento 2 e 1.

Tabela 5.6: Resultados do Experimento 3 sem correferências e entidades.

Modelo	QWK	RMSE
ExtraTreesRegressor	0,399	36,76
CART	0,367	39,19
KNN (k=5)	0,377	35,35

Fonte: (Autor, 2021)

A utilização de modelos regressivos apresentou uma melhora considerável no desempenho da valoração, superando os primeiros experimentos, bem como quase todos os outros valores reportados. Com exceção dos resultados de Fonseca et al., dos quais o melhor resultado

Tabela 5.7: Relação entre FE e FI para estabelecimento da coesão e valoração.

	$FE_i \subset FE_{i+1}$		$FE_i \not\subset FE_{i+1}$	
	Relação	Valor	Relação	Valor
$FI_i \subset FI_{i+1}$	Elaboração	1,00	Mudança de Tópico	0,50
$FI_i \not\subset FI_{i+1}$	Manutenção de tópico	0,75	Mudança de assunto	0,00

Fonte: (Nobre, 2011)

se aproximou tanto dos valores do QWK, quanto do RMSE. Deve-se considerar que a base de redações de Fonseca et al. supera por larga margem a quantidade de instâncias da base de textos construída neste trabalho.

5.3 COESÃO PARA APOIO DE AVALIAÇÕES FORMATIVAS

Conforme descrito no Capítulo 2, a adaptação do método de Nobre (2011) abre um leque de oportunidades de apoio à ações formativas de avaliadores e de educadores, tornando mais rápida a avaliação da qualidade de textos. Nesta seção mostra-se que a valoração da transição entre sentenças - coesão local - e a valoração da relação entre os parágrafos - coesão global - podem ser utilizados como ferramentas de apoio. Tantos os valores de coesão global, quanto local são dados de acordo com a Tabela 5.7.

Durante o cálculo da coesão local, obtêm-se um valor para cada uma dessas transições, considerando os seus respectivos Foco Implícito e Foco Explícito. Assim, é possível compreender como o autor de um texto constrói a estrutura coesiva que sustenta seu texto. Contudo, a análise deste encadeamento é demorado e exige que o avaliador leia o texto, interprete e decodifique as referências e os mecanismos de coesão para obter uma valoração para cada transição entre os pares de sentenças.

O método de valoração da coesão aqui apresentado realiza o cálculo dos valores de coesão local entre os pares de sentenças do texto fornecido, poupando o avaliador desta tarefa, sem contudo dispensar a revisão deste. Assim, o método permite calcular os valores intermediários que compõem o valor da coesão textual e apresentá-los para os avaliadores de forma gráfica.

Por meio da avaliação gráfica, o avaliador pode entender de forma rápida como o texto está estruturado e avaliar tanto a qualidade coesiva, quanto os pontos causadores de problemas dentro do texto. Para ilustrar essa possibilidade de avaliação, na Tabela 5.8 tem-se um texto separado por sentenças e na Figura 5.6 é apresentada a avaliação gráfica deste texto.

A coesão local é avaliada de forma contínua entre os pares subsequentes de sentenças. Assim, cada ponto no gráfico indica uma valoração para a relação coesiva entre um par de sentenças. Também é possível fazer uma avaliação de como a transição entre os parágrafos se dá, pois o ponto de transição da última sentença de um parágrafo para a primeira sentença do próximo está delimitada. Assim, apenas com inspeção visual é possível determinar se há mudança de assuntos reiteradas ou a ausência de sentenças que desenvolvam o tópico proposto em uma redação. Desta forma, é possível valorar de forma rápida o nível de coesão atingido no texto avaliado. Já para avaliar a coesão global deve-se executar outro tipo de inspeção visual.

Da mesma forma que é feito para a coesão local, pode-se ter uma representação gráfica dos relacionamento entre os parágrafos. Contudo, a avaliação dos parágrafos, diferentemente das sentenças, é feita considerando os relacionamento entre todos os pares de parágrafos do texto. Para que dessa forma, seja possível perceber como as entidades e referências se propagam no texto como um todo, constituindo os suportes da coesão global. Na Figura 5.7 temos uma

Tabela 5.8: Texto separado em sentenças.

ID	Sentença
0	Embora as pessoas acreditem que a mulher só obteve seu papel nos últimos séculos, há registros de mulheres que se destacaram na antiguidade clássica, como é o caso da Cleópatra e Hatshepsut (rainhas do Egito), mulheres espartanas e mulheres romanas, apesar dessas não terem uma vida política ativa.
1	Houve progressos e retrocessos ao longo da história nesse âmbito.
2	No entanto, no Brasil, o que se vê é que as mulheres foram tardiamente e gradativamente conquistando seus direitos.
3	A partir de 1827, a mulher brasileira passou a ter brandos direitos, como permissão de estudar, e em 1888 destaca-se a ‘pioneira’ em ‘questões diplomáticas’, Princesa Isabel, que assinou a Lei Áurea, a qual aboliu a escravidão.
4	Apesar de ela ter um papel fundamental na história do Brasil, demorou a outras mulheres se destacarem e, após muitas lutas, Bertha Luz, em 1922, cria a Federação Brasileira pelo Progresso Feminino, que foi o estopim para a conquista do voto feminino.
5	A partir daí, a mulher começou a progredir um pouco mais rapidamente, tendo em 1935 a primeira parlamentar eleita, Carlota de Queiroz.
6	Conforme a constituição, todos são iguais perante a lei não havendo mais distinção de sexo, no entanto ainda há em pleno século XXI preconceitos machistas vindos tanto de homem como de mulheres.
7	Mas isso não é problema para abalar a boa realidade vivida pelas mulheres, já que elas são maioria no país e no eleitorado uma vez que 70 milhões de mulheres votam sendo cinco milhões a mais que os homens, segundo a Agência Nacional.
8	O sexo frágil com certeza já se tornou forte, não só na política, mas nas universidades as mulheres estão ascendendo cada vez mais em relação aos homens, e nas últimas décadas as mulheres está se matriculando 22% a mais que os homens, segundo Instituto Nacional de Estudos e Pesquisas Educacionais (Inep).
9	No entanto nessa maré de rosas, as mulheres sofrem com a discriminação quanto ao salário, pois muitas exercem cargos maiores que de homens e ganham salários inferiores.
10	Mas paulatinamente essas guerreiras vão enfrentar e, com certeza, vencer mais essa batalha!

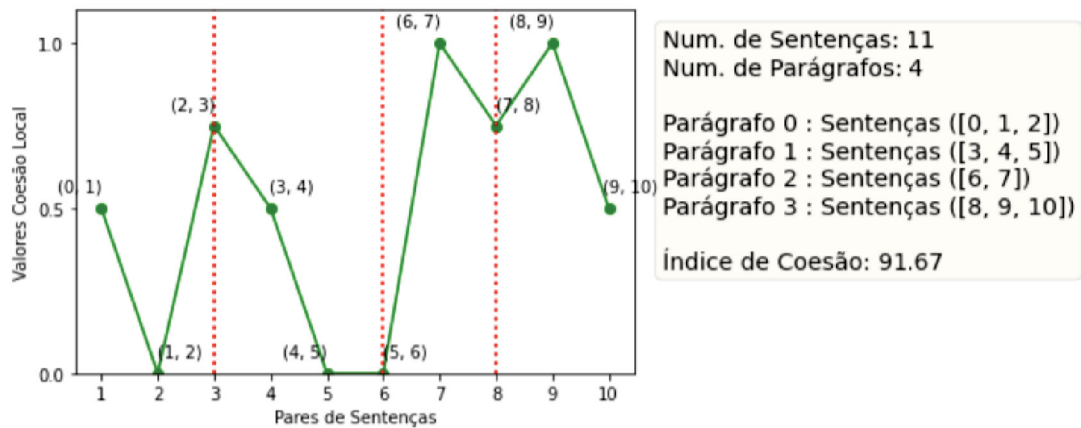
Fonte: (Autor, 2021)

representação destes relacionamentos, onde é possível perceber que cada seção do gráfico traz a informação de como cada parágrafo se relaciona com os demais.

Ao utilizar o gráfico do relacionamento dos pares de parágrafos, o avaliador pode verificar se entre o primeiro e o último parágrafo ocorre um processo de manutenção ou retomada, ao invés, por exemplo, de uma valoração relacionada à mudança de assunto. O que não seria o indicado para um texto bem estruturado, já indicando que entre estes parágrafos há uma deficiência na forma como a coesão global está construída.

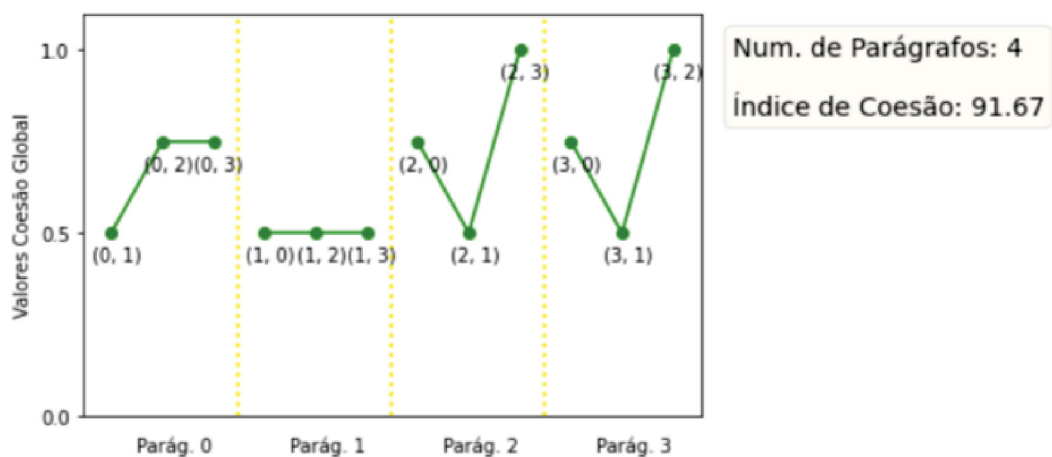
Nas propostas de VAR é comum que seja fornecido uma saída única - geralmente um valor numérico em uma escala - como forma de valorar uma dimensão específica, ou a qualidade geral de um texto. Neste estudo, buscou-se, além da valoração, fornecer, por meio de um método automático, ao avaliador um ganho de tempo, permitindo a este selecionar os textos mais

Figura 5.6: Valores da coesão local dos pares de sentenças calculados pelo método.



Fonte: (Autor, 2021)

Figura 5.7: Valores da coesão global dos pares de parágrafos calculados pelo método.



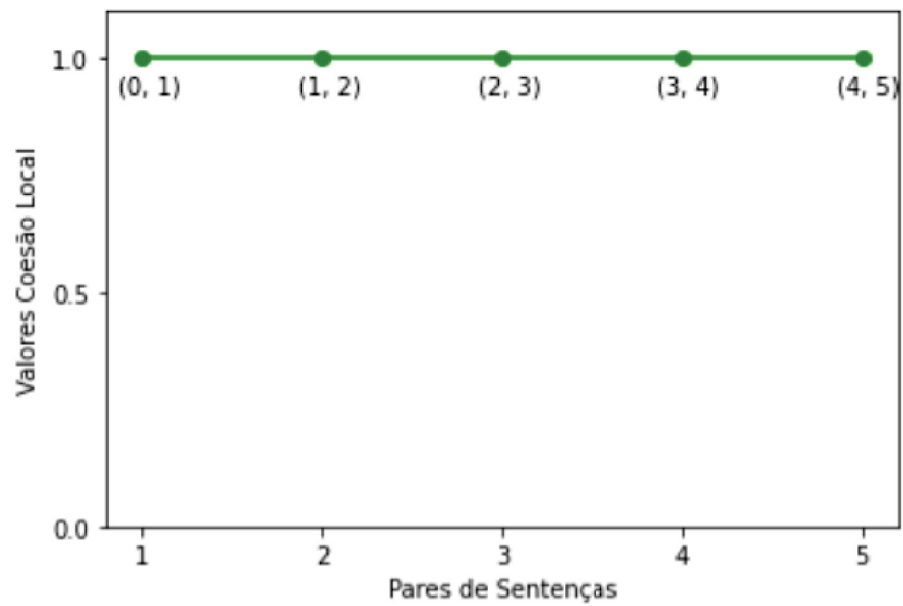
Fonte: (Autor, 2021)

problemáticos em relação à forma ou aqueles que apresentem deficiências mais graves do ponto de vista da qualidade coesiva e abordá-los de forma mais criteriosa.

Avaliando-se os textos por meio das representações gráficas das transições entre suas sentenças e parágrafos é possível identificar problemas de estruturação. Como exemplo, podemos ter um texto onde a transição entre as sentenças ocorra de forma intermitente, conduzindo a um texto onde não ocorre uma gradação na exposição de ideias. Textos deste tipo, fazem com que o leitor não consiga focar no que está sendo transmitido, pois a todo momento ocorrem mudanças de assuntos ou de tópicos.

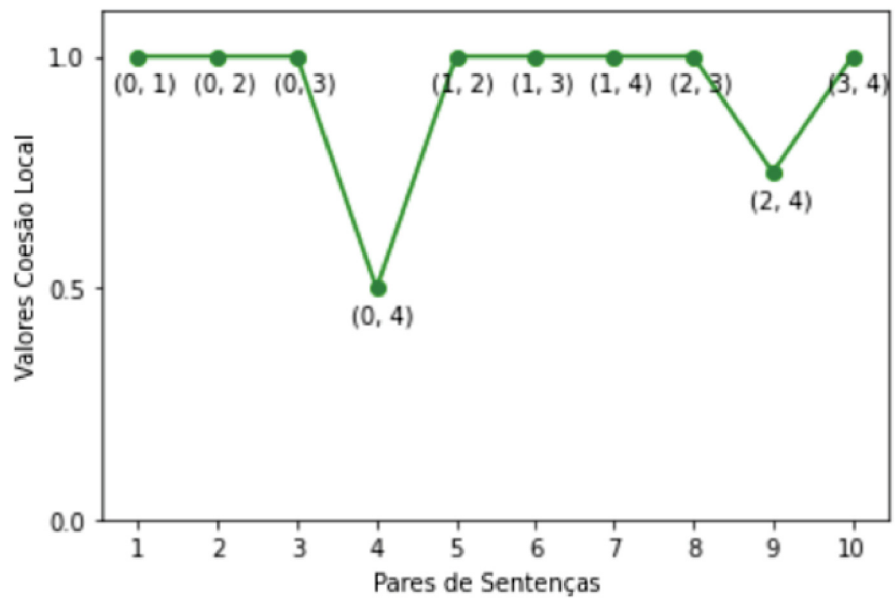
Para exemplificar o uso da proposta de avaliação consideremos as representações gráficas de dois textos apresentados nas Figura <> e Figura <>. O texto da Figura <> tem apenas transições de elaboração entre suas sentenças, não apresentando assim, uma gradação de ideias, nem mesmo nas transições entre parágrafos. Esta ausência de variação no índice de coesão local, já indica que o texto apresenta problemas na sua estruturação. Diferentemente, consideremos a representação apresentada na Figura <>, onde é possível perceber claramente a transição do parágrafo inicial para o segundo e a elaboração no processo de fechamento do texto.

Figura 5.8: Exemplo de texto contendo apenas transições de elaboração.



Fonte: (Autor, 2021)

Figura 5.9: Exemplo de texto com variações nas relações entre sentenças.



Fonte: (Autor, 2021)

Por meio da inspeção visual dos índices de coesão entre sentenças, que pode ser estendida para os parágrafos, é possível identificar de forma rápida o tamanho dos parágrafos e como as ideias e conceitos estão sendo conectadas ao longo de todo o texto. De posse dessa

informação e do conhecimento e experiência do avaliador, o mesmo pode identificar problemas estruturais, ausência de argumentação, exposição de ideias e conceitos.

Com o uso das representações gráficas é possível triar, agrupar e selecionar textos. Também é possível fazer um processo de acompanhamento simultâneo da leitura em conjunto com a análise gráfica. Assim, acredita-se que essa proposta de avaliação apresentada pode contribuir para o desenvolvimento de novas ferramentas de apoio ao desenvolvimento da escrita e auxiliar na correção de textos feita por avaliadores humanos.

5.4 CONCLUSÃO

O Experimento 1 demonstrou que as informações obtidas a partir das transições da sentenças, mensuradas a partir das operações efetuadas sobre as listas de foco explícito e implícito, possuem valor quando consideradas de forma isoladas, ainda que não sejam suficientes para valorar a coesão com um todo.

Os resultados do Experimento 2 mostraram a importância do uso dos métodos de resolução de correferência e da execução dos processos de vinculação de desambiguação de entidades que são feitos em momentos anteriores à descoberta de determinação das cadeias de correferências. Estes resultados não podem ainda ser considerados, quando comparados com outros trabalhos, o que sugere que além da propagação das entidades, seus referentes e as cadeias de correferências, existem outros aspectos que influenciam nos valores finais da coesão que precisam ser capturados pela análise de outros aspectos, como a ortografia, a gramática e os aspectos da coesão projetados na coerência.

O Experimento 3, que traz uma reprodução semelhante ao método de Nobre (2011), mostrou que sem a extração de um largo conjunto de atributos ou uso de métodos de alta complexidade é possível obter resultados que superem trabalhos similares e se aproximem do estado-da-arte, com um esforço computacional e humano, muito menor.

Além dos experimentos, também apresentou-se uma proposta de uso do método para inspeção visual da qualidade coesiva de textos. O que pode permitir o desenvolvimento de ferramentas que auxiliem educadores nos processos de avaliação formativa de aprendizagem da escrita.

Neste capítulo mostrou-se as possibilidades de resultados que o método proposto pode trazer, comparando-o com trabalhos correlatos e na proposição de uma análise visual. O conjunto de experimento ainda mostrou que, mesmo sem utilizar métodos de engenharia de atributos e extração de um largo conjunto de características, nem uso de modelos de aprendizagem complexos é possível obter bons resultados, em comparação com o estado-da-arte, considerando custo e tempo.

6 CONCLUSÃO

Neste capítulo apresenta-se as conclusões, partindo dos resultados obtidos nos Capítulos 4 e 5, e são apresentadas sugestões de trabalhos que futuramente venham a ser realizados tendo este estudo como base.

6.1 CONCLUSÕES GERAIS

Dada a escassez de trabalhos voltados para avaliação de textos em português e que suportem ações educacionais formativas na aprendizagem da escrita, a principal contribuição deste trabalho foi a investigação, análise e desenvolvimento que resultou na adaptação de um método de valoração da coesão textual para o português brasileiro, que se aproxima da valoração do corretor humano. A arquitetura do método criado, integra diferentes propostas para diversas tarefas de PLN, como anotações sintáticas e semânticas, detecção de menção de entidades, resolução de entidade nomeada, vinculação de entidades, resolução de correferências, busca de sinônimos e correção ortográfica e gramatical. A integração destas propostas fornece as informações necessárias para a análise e valoração da coesão, baseado em um método desenvolvido especificamente para o português. Este método, não exige a extração de um largo conjunto de atributos, nem a criação e treinamento de modelos de alta complexidade, mas ainda assim mostra-se viável para processos de avaliação da qualidade coesiva de um texto.

Por meio do protótipo desenvolvido para validar o método, este trabalho também pode contribuir para o desenvolvimento de ferramentas de apoio a avaliadores e educadores em processos formativos de aprendizagem da escrita, podendo prover ganhos de tempo em tarefas de análise de grandes volumes de textos. Este tipo de recurso pode ser de grande valia em um país como o nosso, onde o número de educadores é pequeno se comparado ao número de alunos, especialmente nos anos iniciais da educação básica.

Uma contribuição secundária foi a criação e disponibilização de recursos linguístico-computacionais abertos e disponíveis aqui desenvolvidos, como a criação de uma base de textos anotados com informações morfo-sintáticas e notas de todas as competências textuais exigidas nas correções de redações. Conforme mostrado anteriormente, a quantidade de recursos deste tipo abertos e publicamente disponíveis é baixa. Além das informações obtidas na construção da base de textos, com a execução do protótipo que implementa o modelo proposto, a base conta com os valores de coesão local e coesão global, além dos valores do relacionamento entre as sentenças e os parágrafos do texto. Também destaca-se a criação e treinamento de um modelo de vinculação de entidades para o português brasileiro, baseado em dados da Wikipedia que poderá ser expandido para um modelo de maior abrangência, não ficando restrito apenas ao domínio dos textos dissertativos.

Além dos recursos linguísticos, ainda tem-se uma contribuição com a proposição de uso dos resultados obtidos pelo modelo para gerar gráficos descritores da coesão local e global. O que, de forma mais ampla, permite a sua adoção em sistemas de AAR e sistemas de melhoria da escrita, podendo ser adaptado para outros tipos de textos além daqueles nos quais foi testado. Considerando que o método não apresenta dependência de um corpus ou de dados anotados manualmente, no tocante à informações de discurso. O uso de métodos como este pode trazer perspectivas diferentes no desenvolvimento de sistemas de avaliação automática, pois a avaliação de competências específicas fundamentada em teorias linguísticas pode gerar

avaliadores automáticos mais velozes, robustos e eficazes de avaliação textual, sem que este seja apenas um contador de letras, pontuações ou quaisquer outros elementos superficiais de um texto.

6.2 TRABALHOS FUTUROS

Considerando os resultados obtidos, sugere-se:

1. Testar o método com diferentes bases de texto;
2. Integrar o método de valoração com métodos de extração de atributos;
3. Integrar o método de valoração com métodos de redes neurais na avaliação de outras competências correlacionadas à coesão como a coerência;
4. Utilizar a valoração de coesão com outros métodos de AAR;
5. Desenvolver novas ferramentas que permitam identificar melhor as referências anafóricas e as cadeias de correferência;
6. Transformar o protótipo em uma ferramenta de apoio à escrita e desta forma validar a proposição de uso do método como suporte de processos formativos e avaliar o ganho oferecido.

REFERÊNCIAS

- Amorim, E. e Veloso, A. (2017). A multi-aspect analysis of automatic essay scoring for brazilian portuguese. Em *Proceedings of the Student Research Workshop at the 15th Conference of the European Chapter of the Association for Computational Linguistics*, páginas 94–102.
- Bick, E. (2014). Palavras, a constraint grammarbased parsing system for portuguese. *Working with Portuguese corpora*, páginas 279–302.
- Bittencourt Júnior, J. A. S. et al. (2020). Avaliação automática de redação em língua portuguesa empregando redes neurais profundas.
- Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., Becker, C., Cyganiak, R. e Hellmann, S. (2009). Dbpedia-a crystallization point for the web of data. *Journal of web semantics*, 7(3):154–165.
- CHARROLES, M. (1988). Introdução aos problemas da coerência dos textos: abordagem teórica e estudo das práticas pedagógicas. *O texto: leitura e escrita*. Campinas, SP: Pontes, páginas 39–85.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46.
- Crossley, S. A., Kyle, K. e McNamara, D. S. (2016). The tool for the automatic analysis of text cohesion (taaco): Automatic assessment of local, global, and text cohesion. *Behavior research methods*, 48(4):1227–1237.
- de Souza, V. M. A. e Feltrim, V. D. (2013). A coherence analysis module for scipo: providing suggestions for scientific abstracts written in portuguese. *Journal of the Brazilian Computer Society*, 19(1):59–73.
- Dias, M. e Pardo, T. (2015). A discursive grid approach to model local coherence in multi-document summaries. Em *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, páginas 60–67.
- Dias, M. d. S. (2016). *Investigação de modelos de coerência local para sumários multidocumento*. Tese de doutorado, Universidade de São Paulo.
- Dikli, S. (2006). An overview of automated scoring of essays. *The Journal of Technology, Learning and Assessment*, 5(1).
- Feltrim, V. D., Teufel, S., das Nunes, M. G. V. e Aluísio, S. M. (2006). Argumentative zoning applied to critiquing novices' scientific abstracts. Em *Computing Attitude and Affect in Text: Theory and Applications*, páginas 233–246. Springer.
- Fonseca, E., Medeiros, I., Kamikawachi, D. e Bokan, A. (2018a). Automatically grading brazilian student essays. Em *International Conference on Computational Processing of the Portuguese Language*, páginas 170–179. Springer.
- Fonseca, E. B. et al. (2018b). Resolução de correferência nominal usando semântica em língua portuguesa.

- Grosz, B. J. e Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational linguistics*, 12(3):175–204.
- Honnibal, M. e Montani, I. (2017). spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. To appear.
- Ji, H. e Grishman, R. (2011). Knowledge base population: Successful approaches and challenges. páginas 1148–1158.
- Jorge, M. L. d. R. C., Dias, M. S. e Pardo, T. A. S. (2014). Building a language model for local coherence in multi-document summaries using a discourse-enriched entity-based model. Em *2014 Brazilian Conference on Intelligent Systems*, páginas 44–49. IEEE.
- Jurafsky, D. (2000). *Speech & language processing*. Pearson Education India.
- Jurafsky, D. e Martin, J. H. (2019). *Speech and Language Processing (3rd draft ed.)*. Stanford Univ.
- Ke, Z. e Ng, V. (2019). Automated essay scoring: A survey of the state of the art. Em *IJCAI*, volume 19, páginas 6300–6308.
- Kinoshita, J., Salvador, L. N., Menezes, C. E. e Silva, W. (2007). Cogroo-an openoffice grammar checker. Em *Seventh International Conference on Intelligent Systems Design and Applications (ISDA 2007)*, páginas 525–530. IEEE.
- Koch, I. G. V. (1999). *A coesão textual*. Editora Contexto.
- Koch, I. G. V. e Travaglia, L. C. (1989). *Texto e coerência*, volume 4. Cortez Editora.
- Koch, I. G. V. e Travaglia, L. C. (1990). *A coerência textual*. Contexto São Paulo.
- Lai, A. e Tetreault, J. (2018). Discourse coherence in the wild: A dataset, evaluation and methods. *arXiv preprint arXiv:1805.04993*.
- Li, J. e Jurafsky, D. (2016). Neural net models for open-domain discourse coherence. *arXiv preprint arXiv:1606.01545*.
- Lima, L. M. (2017). *Linguística Textual*. Editora e Distribuidora Educacional S.As.
- Lin, Z., Ng, H. T. e Kan, M.-Y. (2011). Automatically evaluating text coherence using discourse relations. Em *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, páginas 997–1006.
- Marinho, J. C., Anchieta, R. T. e Moura, R. S. (2021). Essay-br: a brazilian corpus of essays. *arXiv preprint arXiv:2105.09081*.
- Maziero, E. G., Pardo, T. A., Di Felippo, A. e Dias-da Silva, B. C. (2008). A base de dados lexical ea interface web do tep 2.0: thesaurus eletrônico para o português do brasil. Em *Companion Proceedings of the XIV Brazilian Symposium on Multimedia and the Web*, páginas 390–392.
- Mendes, P. N., Jakob, M., García-Silva, A. e Bizer, C. (2011). Dbpedia spotlight: shedding light on the web of documents. Em *Proceedings of the 7th international conference on semantic systems*, páginas 1–8.

- Mihalcea, R. e Csomai, A. (2007). Wikify! linking documents to encyclopedic knowledge. Em *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*, páginas 233–242.
- Nguyen, D. T. e Joty, S. (2017). A neural local coherence model. Em *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, páginas 1320–1330.
- Nobre, J. C. S. (2011). *Modelo computacional para valoração e avaliação de redações baseado em lógica Fuzzy*. Tese de doutorado, Instituto Tecnológico de Aeronáutica, São José dos Campos, São Paulo, Brazil.
- Nobre, J. C. S. e Pellegrino, S. R. M. (2010). Anac: um analisador automático de coesão textual em redação. Em *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, volume 1.
- Page, E. B. (1966). Project essay grade: Peg.
- Perelman, L. (2014). When “the state of the art” is counting words. *Assessing Writing*, 21:104–111.
- Santos, C. M. e Kroeff, R. F. d. S. (2018). A contribuição do feedback no processo de avaliação formativa.
- Shermis, M. D. (2014). State-of-the-art automated essay scoring: Competition, results, and future directions from a united states demonstration. *Assessing Writing*, 20:53–76.
- Sidner, C. L. (1979). Towards a computational theory of definite anaphora comprehension in english discourse. Relatório técnico, Massachusetts Inst of Tech Cambridge Artificial Intelligence lab.
- Silva, L. L. d. (2016). Análise automática de coerência textual em resumos científicos: avaliando quebras de linearidade. Dissertação de Mestrado, Universidade Estadual de Maringá.
- Silva, W. D. C. M. e Finger, M. (2013). Improving cogroo: the brazilian portuguese grammar checker. Em *Proceedings of the 9th Brazilian Symposium in Information and Human Language Technology*.
- Souza, V. M. A. d. (2011). Análise automática de coerência semântica em recursos acadêmicos escritos em português. Dissertação de Mestrado, Universidade Estadual de Maringá.
- Taboada, M. e Wieseemann, L. (2010). Subjects and topics in conversation. *Journal of Pragmatics*, 42(7):1816–1828.
- Van Dijk, T. A., Kintsch, W. et al. (1983). Strategies of discourse comprehension.
- Xu, P., Saghir, H., Kang, J. S., Long, T., Bose, A. J., Cao, Y. e Cheung, J. C. K. (2019). A cross-domain transferable neural coherence model. *arXiv preprint arXiv:1905.11912*.

APÊNDICE A – ERROS GRAMATICAIS CONSIDERADOS

Tabela A.1: Tabela de Erros Gramaticais.

Código	Descrição
xml:1	Não ocorre crase antes de palavras masculinas.
xml:2	Não ocorre crase antes de palavras masculinas.
xml:3	Não há crase neste caso, somente no plural ("às").
xml:4	Não acontece crase antes de verbo.
xml:5	Ocorre crase em expressões indicativas de horas.
xml:6	Ocorre crase em expressões indicativas de horas.
xml:9	"Em relação" rege a preposição "a", logo há crase aqui.
xml:10	"Em relação" rege a preposição "a", logo há crase aqui.
xml:11	"Com relação" rege a preposição "a", logo há crase aqui.
xml:12	"Com relação" rege a preposição "a", logo há crase aqui.
xml:13	"Devido" rege a preposição "a", logo há crase aqui.
xml:14	"Devido" rege a preposição "a", logo há crase aqui.
xml:17	O adjetivo concorda com o substantivo a que se refere.
xml:21	O adjetivo concorda com o substantivo a que se refere.
xml:39	A palavra "meio" usada no sentido de "um pouco" (advérbio) é invariável.
xml:40	"Meio" usado no sentido de "metade" (numeral) concorda com o substantivo.
xml:42	"Fazer", quando indica tempo, deve permanecer no singular.
xml:46	Opte pelo verbo "haver", no singular, para indicar tempo decorrido.
xml:47	Opte pelo verbo "haver", no singular, para indicar tempo decorrido.
xml:50	"Haver" no sentido de existir é usado na 3a. pessoa do singular.
xml:51	"Haver" (existir), precedido de verbo aux., é usado na 3a. do singular.
xml:52	Use "eu" ao invés de "mim" como sujeito de verbo no infinitivo.
xml:57	O contrário de "bem" é "mal", e o de "bom" é "mau".
xml:58	O contrário de "bem" é "mal", e o de "bom" é "mau".
xml:61	Palavras negativas atraem o pronome átono para antes do verbo.
xml:65	Pron. relativos e conj. subordinativas atraem o pronome para antes do verbo.
xml:67	Certos advérbios atraem o pronome para antes do verbo.
xml:71	Alguns pronomes indefinidos atraem o pronome para antes do verbo.
xml:73	"só, ou, ora e quer" atraem o pronome para antes do verbo.
xml:75	Conjugação de um verbo irregular no futuro do subjuntivo.
xml:78	"Ir" constrói-se com prep. "a". Há crase com compl. feminino.
xml:79	"Aderir" constrói-se com prep. "a". Há crase com compl. feminino.
xml:80	"Pentencer" constrói-se com prep. "a". Há crase com compl. feminino.
xml:82	"(Des)obedecer" constrói-se com prep. "a". Há crase com compl. feminino.
xml:84	Alguns nomes regem a preposição "a", logo há crase aqui.
xml:86	"(Des)obedecer" constrói-se com preposição "a".
xml:87	Não há crase porque alguns pronomes pessoais não admitem artigo.
xml:88	"Eu" não deve ser preposicionado. Use "a mim".
xml:91	Alguns nomes regem a preposição "a", logo há crase aqui.
xml:94	"Equivaler" constrói-se com prep. "a". Há crase com compl. feminino.

Continuação da Tabela A.1	
Código	Descrição
xml:95	Os artigos concordam com o substantivo a que se referem.
xml:96	Os verbos "evitar" e "usufruir" não regem preposição "de".
xml:97	Regência do verbo incorreta.
xml:103	Os artigos concordam com o substantivo a que se referem.
xml:104	Os artigos concordam com o substantivo a que se referem.
xml:105	Os artigos concordam com o substantivo a que se referem.
xml:106	A expressão correta é "à medida que".
xml:107	É inadequado o uso da preposição "em".
xml:110	"Valorização" rege preposição "de" e não preposição "a".
xml:111	A expressão "ou seja" deve ser isolada por vírgulas.
xml:112	A expressão "ou seja" deve ser isolada por vírgulas.
xml:113	Deve haver vírgula antes de "no entanto".
xml:114	Os determinantes concordam com o substantivo a que se referem.
xml:115	Os numerais concordam com o substantivo a que se referem.
xml:117	Verificou-se erro de concordância entre o sujeito e o verbo.
xml:118	Verificou-se erro de concordância entre o sujeito e o verbo.
xml:121	Verifique o excesso de verbos em sequência.
xml:122	"Atrás" é redundante, devido ao uso do verbo "haver".
xml:124	O adjetivo na função de predicativo concorda com o sujeito.
xml:127	O adjetivo na função de predicativo concorda com o sujeito.
xml:128	O adjetivo na função de predicativo concorda com o verbo.
xml:129	Verificou-se erro de concordância entre o sujeito e o verbo.

Fonte: (Kinoshita et al., 2007)

Tabela A.2: Erros de paronímia, regência e repetição.

Código	Descrição
repetition:DUPLICATED_TOKEN	Repetição de palavras.
government:GOVERNMENT	Regência verbal.
probs:paronyms	Possível confusão entre bebê e bebe.
probs:paronyms	Possível confusão entre porem e porém.
probs:paronyms	Possível confusão entre diagnostico e diagnóstico.
probs:paronyms	Possível confusão entre exercito e exército.
probs:paronyms	Possível confusão entre habito e hábito.
probs:paronyms	Possível confusão entre influencia e influência.
probs:paronyms	Possível confusão entre acumulo e acúmulo.
probs:paronyms	Possível confusão entre critico e crítico.
probs:paronyms	Possível confusão entre Porem e Porém.
probs:paronyms	Possível confusão entre duvida e dúvida.
probs:paronyms	Possível confusão entre divida e dívida.
probs:paronyms	Possível confusão entre publico e público.
probs:paronyms	Possível confusão entre influência e influencia.

Fonte: (Kinoshita et al., 2007)

APÊNDICE B – TEXTOS COM ANOTAÇÕES E DADOS

B.1 SENTENÇAS, LISTAS DE FOCO E ENTIDADES

Sentença: 0

Texto Original: Embora as pessoas acreditem que a mulher só obteve seu papel nos últimos séculos, há registros de mulheres que se destacaram na antiguidade clássica, como é o caso da Cleópatra e Hatshepsut (rainhas do Egito), mulheres espartanas e mulheres romanas, apesar dessas não terem uma vida política ativa.

Lista de Foco Explícito: ['rainhas', 'Egito', 'antiguidade clássica', 'séculos', 'se', 'pessoas', 'caso', 'vida', 'registros', 'Cleópatra', 'Hatshepsut', 'papel', 'mulheres', 'dessas', 'que']

Lista Intermediária de FE: 'caso': 'ocorrência', 'evento', 'conto', 'acontecimento', 'aventura', 'estória', 'envolvimento', 'episódio', 'namoro', 'romance', 'fato', 'vida': 'sal', 'brilho', 'expressão', 'alma', 'existência', 'animação', 'subsistência', 'viveza', 'dias', 'biografia', 'tudo', 'sangue', 'vivacidade', 'viver', 'colorido', 'vivência', 'papel': 'função', 'finalidade', 'que': 'quão'

Lista de Entidades Nomeadas: [Cleópatra, Hatshepsut, Egito]

Lista de Menções DBPedia: ['antiguidade clássica', 'Cleópatra', 'Hatshepsut', 'Egito']

Sentença: 1

Texto Original: Houve progressos e retrocessos ao longo da história nesse âmbito. Lista de Foco Explícito: ['longo', 'retrocessos', 'âmbito', 'progressos', 'história'] Lista Intermediária de FE: 'longo': 'copioso', 'comprido', 'extenso', 'demorado', 'largo', 'dilatado', 'estirado', 'prolixo', 'grande', 'âmbito': 'esfera', 'área', 'ambiente', 'atmosfera', 'circunferência', 'contorno', 'circuito', 'domínio', 'campo', 'periferia', 'história': 'lampana', 'conto', 'peta', 'tradição', 'tretas', 'prego', 'potoca', 'narrativa', 'mentira', 'narração', 'conversa', 'pulha', 'patranha', 'lorota', 'romance'

Lista de Entidades Nomeadas: []

Lista de Menções DBPedia: []

Sentença: 2

Texto Original: No entanto, no Brasil, o que se vê é que as mulheres foram tardiamente e gradativamente conquistando seus direitos.

de Foco Explícito: ['direitos', 'se', 'mulheres', 'que', 'entanto', 'Brasil'] Intermediária de FE: 'que': 'quão', 'entanto': 'entretanto', 'nesteínterim', 'entrementes', 'nestemeiotempo' de Entidades Nomeadas: [Brasil] de Menções DBPedia: ['Brasil']

Sentença: 3

Texto Original: A partir de 1827, a mulher brasileira passou a ter brandos direitos, como permissão de estudar, e em 1888 destaca-se a 'pioneira' em 'questões diplomáticas', Princesa Isabel, que assinou a Lei Áurea, a qual aboliu a escravidão.

Lista de Foco Explícito: ['direitos', 'escravidão', 'mulher', 'partir', 'Princesa Isabel', 'permissão', 'questões', 'Lei Áurea', 'qual']

Lista Intermediária de FE: 'mulher': 'concubina', 'esposa', 'varoa', 'barregã', 'dona', 'amásia', 'amiga', 'fêmea', 'amante', 'partir': 'descender', 'trincar', 'voltar-se', 'despedir-se', 'escaqueirar', 'derivar', 'provir', 'compartilhar', 'crescer', 'rascar', 'desfazer', 'seccionar', 'falhar', 'zarpar',

'partilhar', 'arremeter', 'retirar-se', 'agredir', 'dividir', 'cindir', 'resultar', 'romper', 'atacar', 'despedaçar', 'hastilhar', 'nascer', 'derramar', 'espedaçar', 'torar', 'fender', 'seccionar', 'atroçar', 'carregar', 'fracionar', 'ir-se', 'separar', 'rachar', 'departir', 'acometer', 'remeter', 'fragmentar', 'gretar', 'fissurar', 'deixar', 'vir', 'estroçar', 'repartir', 'espostear', 'saltear', 'sair', 'distribuir', 'compartir', 'largar', 'assaltar', 'atirar-se', 'proceder', 'escacar', 'esfachear', 'opugnar', 'lascar', 'avançar', 'estilhar', 'abalar', 'desaportar', 'estilhaçar', 'fraturar', 'retalhar', 'investir', 'arrancar', 'escavacar', 'ausentar-se', 'aferrar', 'estriçar', 'ser', 'escachar', 'destroçar', 'escindir', 'esfazer', 'ir', 'atorar', 'render', 'esbandalhar', 'espatifar', 'assaltear', 'desbandalhar', 'quebrar', 'permissão': 'faculdade', 'vênia', 'concessão', 'autoridade', 'liberdade', 'consentimento', 'autorização', 'licença'

Lista de Entidades Nomeadas: [Princesa Isabel, Lei Áurea]

Lista de Menções DBPedia: ['Princesa Isabel', 'Lei Áurea', 'escravidão']

Sentença: 4

Texto Original: Apesar de ela ter um papel fundamental na história do Brasil, demorou a outras mulheres se destacarem e, após muitas lutas, Bertha Luz, em 1922, cria a Federação Brasileira pelo Progresso Feminino, que foi o estopim para a conquista do voto feminino.

Lista de Foco Explícito: ['estopim', 'ela', 'voto feminino', 'Bertha Luz', 'se', 'conquista', 'papel', 'mulheres', 'lutas', 'Progresso Feminino', 'história', 'que', 'Federação Brasileira']

Lista Intermediária de FE: 'conquista': 'logro', 'consequimento', 'colheita', 'consecução', 'triunfo', 'vindima', 'adquirição', 'aquisição', 'granjeio', 'vencimento', 'messe', 'tomada', 'alcance', 'absorção', 'alcançamento', 'obtenção', 'adquisição', 'adquirimento', 'vitória', 'troféu', 'sedução', 'anexação', 'ganho', 'papel': 'função', 'finalidade', 'história': 'lampana', 'conto', 'peta', 'tradição', 'tretas', 'prego', 'potoca', 'narrativa', 'mentira', 'narração', 'conversa', 'pulha', 'patranha', 'lorota', 'romance', 'que': 'quão'

Lista de Entidades Nomeadas: [Brasil, Bertha Luz, Federação Brasileira, Progresso Feminino]

Lista de Menções DBPedia: ['Brasil', 'voto feminino']

Sentença: 5

Texto Original: A partir daí, a mulher começou a progredir um pouco mais rapidamente, tendo em 1935 a primeira parlamentar eleita, Carlota de Queiroz.

Lista de Foco Explícito: ['Carlota de Queiroz', 'mulher', 'parlamentar']

Lista Intermediária de FE: 'mulher': 'concubina', 'esposa', 'varoa', 'barregã', 'dona', 'amásia', 'amiga', 'fêmea', 'amante', 'parlamentar': 'conferenciar', 'parlamentear'

Lista de Entidades Nomeadas: [Carlota de Queiroz]

Lista de Menções DBPedia: ['Carlota']

Sentença: 6

Texto Original: Conforme a constituição, todos são iguais perante a lei não havendo mais distinção de sexo, no entanto ainda há em pleno século XXI preconceitos machistas vindos tanto de homem como de mulheres.

Lista de Foco Explícito: ['todos', 'lei', 'constituição', 'homem', 'sexo', 'mulheres', 'distinção', 'século XXI', 'entanto', 'preconceitos']

Lista Intermediária de FE: 'lei': 'regra', 'constituição', 'preceito', 'norma', 'estatuto', 'prescrição', 'ordem', 'código', 'constituição': 'composição', 'anatomia', 'lei', 'compleição', 'carnadura', 'organização', 'corporatura', 'formação', 'figura', 'presença', 'caráter', 'aspecto', 'corpo', 'natureza', 'vulto', 'estatuto', 'substância', 'essência', 'homem': 'homem-feito', 'serhumano', 'varão', 'macho', 'marido', 'humanidade', 'esposo', 'distinção': 'esbelteza', 'marcação', 'excelência', 'ponta', 'qualidade', 'superioridade', 'sublimidade', 'donaire', 'diferenciação', 'predicamento',

'honraria', 'galhardia', 'verniz', 'grandeza', 'honras', 'elevação', 'separação', 'evidência', 'ordem', 'airosidade', 'plana', 'classe', 'nobreza', 'altura', 'discernimento', 'voga', 'categoria', 'projeção', 'elegância', 'esbeltez', 'notabilidade', 'relevo', 'caracterização', 'mérito', 'grandiosidade', 'destaque', 'glória', 'garbo', 'diferença', 'majestade', 'conspicuidade', 'realce', 'apercepção', 'discriminação', 'torneio', 'grau', 'entanto': 'entretanto', 'nesteínterim', 'entrementes', 'nestemeio-tempo'

Lista de Entidades Nomeadas: []

Lista de Menções DBPedia: ['século XXI']

Sentença: 7

Texto Original: Mas isso não é problema para abalar a boa realidade vivida pelas mulheres, já que elas são maioria no país e no eleitorado uma vez que 70 milhões de mulheres votam sendo cinco milhões a mais que os homens, segundo a Agência Nacional.

Lista de Foco Explícito: ['país', 'Agência Nacional', 'elas', 'isso', 'mulheres', 'eleitorado', 'homens', 'problema', 'realidade', 'maioria', 'vez']

Lista Intermediária de FE: 'país': 'nação', 'território', 'problema': 'quesito', 'questão', 'complicação', 'realidade': 'efetividade', 'exatidão', 'veras', 'real', 'verdade', 'maioria': 'grosso', 'geral', 'vez': 'viravolta', 'lugar', 'oportunidade', 'tempo', 'hora', 'ensejo', 'dose', 'vicissitude', 'margem', 'porção', 'alternativa', 'turno'

Lista de Entidades Nomeadas: [Agência Nacional]

Lista de Menções DBPedia: []

Sentença: 8

Texto Original: O sexo frágil com certeza já se tornou forte, não só na política, mas nas universidades as mulheres estão ascendendo cada vez mais em relação aos homens, e nas últimas décadas as mulheres está se matriculando 22% a mais que os homens, segundo Instituto Nacional de Estudos e Pesquisas Educacionais (Inep).

Lista de Foco Explícito: ['política', 'sexo frágil', 'Inep', 'Instituto Nacional de Estudos e Pesquisas Educacionais', 'décadas', 'mulheres', 'homens', 'mais', 'relação', 'certeza', 'universidades']

Lista Intermediária de FE: 'política': 'educação', 'astúcia', 'cortesia', 'ardileza', 'civildade', 'trato', 'amenidade', 'polidez', 'manha', 'urbanidade', 'mais': 'restante', 'outravez', 'resto', 'denovo', 'antes', 'preferentemente', 'além disso', 'compreensão', 'também', 'relação': 'lista', 'conexão', 'relacionamento', 'listagem', 'rol', 'referência', 'encadeamento', 'analogia', 'abraço', 'subordinação', 'vinculação', 'vínculo', 'mapa', 'aproximação', 'respeito', 'enumeração', 'quadro', 'catálogo', 'encadeação', 'concernência', 'afinidade', 'nomenclatura', 'união', 'elo', 'ligação', 'certeza': 'seguridade', 'persuasão', 'convencimento', 'persuadimento', 'segurança', 'convicção', 'firmeza', 'segurança'

Lista de Entidades Nomeadas: [Instituto Nacional de Estudos e Pesquisas Educacionais, Inep]

Lista de Menções DBPedia: ['sexo frágil', 'Inep']

Sentença: 9

Texto Original: No entanto nessa maré de rosas, as mulheres sofrem com a discriminação quanto ao salário, pois muitas exercem cargos maiores que de homens e ganham salários inferiores.

Lista de Foco Explícito: ['salários', 'maré', 'rosas', 'cargos', 'muitas', 'mulheres', 'homens', 'que', 'discriminação', 'entanto']

Lista Intermediária de FE: 'que': 'quão', 'discriminação': 'discernimento', 'apercepção', 'parcialidade', 'distinção', 'entanto': 'entretanto', 'neste ínterim', 'entrementes', 'neste meio tempo'

Lista de Entidades Nomeadas: []

Lista de Menções DBPedia: ['maré']

Sentença: 10

Texto Original: Mas paulatinamente essas guerreiras vão enfrentar e, com certeza, vencer mais essa batalha!

Lista de Foco Explícito: ['batalha', 'guerreiras', 'certeza']

Lista Intermediária de FE: 'batalha': 'certame', 'embate', 'esforço', 'discussão', 'empenho', 'contenda', 'guerra', 'certâmen', 'duelo', 'recontro', 'controvérsia', 'ação', 'encontro', 'justa', 'luta', 'combate', 'choque', 'conflito', 'peleja', 'torneio', 'certeza': 'seguridade', 'persuasão', 'convencimento', 'persuadimento', 'segureza', 'convicção', 'firmeza', 'segurança'

Lista de Entidades Nomeadas: []

Lista de Menções DBPedia: []

B.2 CADEIAS DE CORREFERÊNCIA

Tabela B.1: Tabela de Cadeias de Correferências.

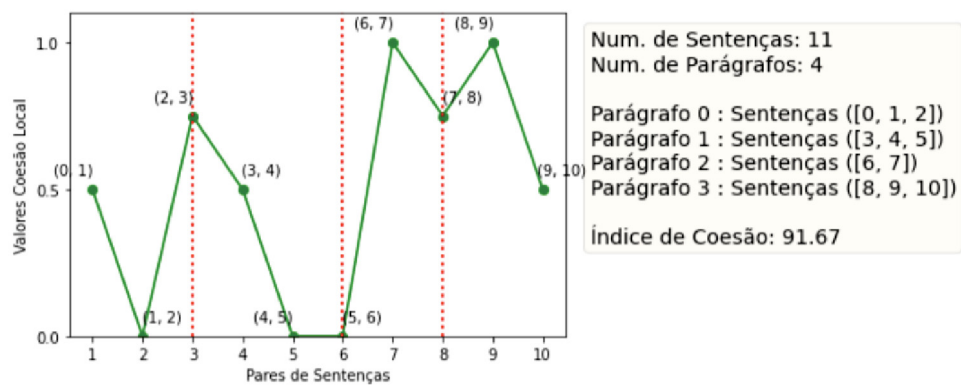
Número da Cadeia	Correferências	
	Sentença	Sintagma
Cadeia 1	0	as pessoas
	0	a mulher
	5	a mulher
	6	homem
	7	os homens
	8	os homens
	9	homens
Cadeia 2	0	mulheres
	2	as mulheres
	6	mulheres
	7	as mulheres
	7	mulheres
	8	as mulheres
	9	as mulheres
Cadeia 3	0	o caso de a Cleópatra e Hatshepsut
	1	a história
	4	a história
Cadeia 4	0	Hatshepsut
	0	rainhas de o Egito
Cadeia 5	2	o Brasil
	4	o Brasil
Cadeia 6	4	ela
	4	um papel fundamental
Cadeia 7	6	século XXI preconceitos machistas

Continuação da Tabela B.1		
Número da Cadeia	Correferências	
	Sentença	Sintagma
	7	maioria
Cadeia 8	8	certeza
	10	certeza
Cadeia 9	8	Pesquisas Educacionais
	8	Inep

Fonte: (Autor, 2021)

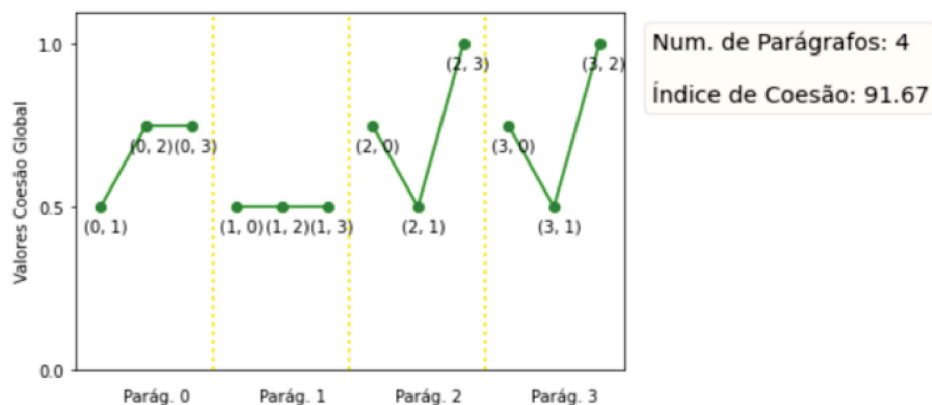
B.3 COESÃO LOCAL E GLOBAL

Figura B.1: Valores da coesão local dos pares de sentenças.



Fonte: (Autor, 2021)

Figura B.2: Valores da coesão local dos pares de parágrafos.



Fonte: (Autor, 2021)

B.4 ERROS DETECTADOS PELO COGROO

space:EXTRA_BETWEEN_WORDS - Excesso de espaços entre as palavras. xml:124 - O adjetivo na função de predicativo concorda com o sujeito.

B.5 ENTIDADES, ALIAS E CANDIDATOS

```
{
  "séculos": {
    "label": 16544881809274863474,
    "kb_id": "Q578",
    "candidates": ["século", "Século", "Século I", "Século X"]},
  "Cleópatra": {
    "label": 4317129024397789502,
    "kb_id": "Q635",
    "candidates": ["Cleópatra VII", "Cleópatra Berenice"]},
  "Hatshepsut": {
    "label": 4317129024397789502,
    "kb_id": "",
    "candidates": []},
  "Egito": {
    "label": 385,
    "kb_id": "Q79",
    "candidates": ["Libra do Egito", "Antigo Egito"]},
  "Brasil": {
    "label": 385,
    "kb_id": "Q155",
    "candidates": ["GE Brasil", "Rede Brasil", "Arte do Brasil"]},
  "brasileira": {
    "label": 16544881809274863474,
    "kb_id": "Q873625",
    "candidates": ["A Brasileira", "Brasileiras", "Arte brasileira"]},
  "Princesa Isabel": {
    "label": 4317129024397789502,
    "kb_id": "Q464449",
    "candidates": ["Princesa Isabella", "Princesa Isabel do Brasil"]},
  "Lei Áurea": {
    "label": 7654241940133152407,
    "kb_id": "Q1519167",
    "candidates": []},
  "Bertha Luz": {
    "label": 4317129024397789502,
    "kb_id": "",
    "candidates": []},
  "Federação Brasileira": {
    "label": 383,
    "kb_id": "Q43702",
    "candidates": ["Federação", "Confederação Brasileira de Futebol"]},
  "Progresso Feminino": {
    "label": 383,
    "kb_id": "",
    "candidates": []},
  "Carlota de Queiroz": {
    "label": 4317129024397789502,
    "kb_id": "",
    "candidates": []},
  "século": {
    "label": 16544881809274863474,
    "kb_id": "Q578",
    "candidates": ["Século I", "Século X", "séculos"]},
}
```

```

"preconceitos": {"label": 16544881809274863474,
                  "kb_id": "Q179742",
                  "candidates": ["Preconceito", "Preconceito social",
                                "Preconceito sexual"]},
"Agência Nacional": {"label": 383,
                      "kb_id": "Q16144331",
                      "candidates": ["Agência Nacional de Águas",
                                    "Agência de Segurança Nacional"]},
"décadas": {"label": 16544881809274863474,
             "kb_id": "Q39911",
             "candidates": ["Década", "década", "Década de vinte"]},
"Instituto Nacional de Estudos e Pesquisas Educacionais": {
    "label": 383,
    "kb_id": "Q6041403",
    "candidates": ["Instituto
                   Nacional
                   de Estudos e
                   Pesquisas Educacionais
                   Anísio Teixeira"]},
"Inep": {"label": 383, "kb_id": "Q6041403", "candidates": []}}

```

B.6 ANOTAÇÕES PARSE PALAVRAS

```

Embora [embora] <clb> <clb-fs> <*> KS @SUB #1->4
as [o] <artd> DET F P @>N #2->3
pessoas [pessoa] <H> N F P @SUBJ> #3->4
acreditem [acreditar] <vH> <mv> V PR 3P SUBJ VFIN
@FS-ADVL> #4->17
que [que] <clb> <clb-fs> KS @SUB #5->9
a [o] <artd> DET F S @>N #6->7
mulher [mulher] <Hbio> <Hfam> <Hattr> N F S @SUBJ> #7->9
só [só] <setop> ADV @ADVL> #8->9
obteve [obter] <vH> <mv> V PS 3S IND VFIN @FS-<ACC #9->4
seu [seu] <poss 3S> <si> DET M S @>N #10->11
papel [papel] <mat> <cc-r> <ac> N M S @<ACC #11->9
em [em] <sam-> PRP @<ADVL #12->9
os [o] <artd> <-sam> DET M P @>N #13->15
últimos [último] <NUM-ord> ADJ M P @>N #14->15
séculos [século] <dur> <per> N M P @P< #15->12
$, #16->0
há [haver] <fmc> <mv> V PR 3S IND VFIN @FS-STA #17->0
registros [registro] <cc-r> N M P @<ACC #18->17
de [de] <np-close> PRP @N< #19->18
mulheres [mulher] <Hbio> <Hfam> <Hattr> N F P @P< #20->19
que [que] <clb> <clb-fs> <rel> SPEC M S @SUBJ> #21->23
se [se] <refl> PERS M 3P ACC @ACC> #22->23

```

destacaram [destacar] <mv> <v-pron> <np-close> V PS/MQP 3P IND
 VFIN @FS-N< #23->20
 em [em] <sam-> PRP @<ADVL #24->23
 a [o] <-sam> <artd> DET F S @>N #25->26
 antiguidade [antiguidade] <per> <cc> N F S @P< #26->24
 clássica [clássico] <nh> <np-close> ADJ F S @N< #27->26
 \$, #28->0
 como [como] <clb> <clb-fs> <rel> <ks> ADV @ADVL> #29->30
 é [ser] <mv> V PR 3S IND VFIN @FS-<ADVL #30->23
 o [o] <artd> DET M S @>N #31->32
 caso [caso] <event> <ac-cat> <sem> N M S @<SUBJ #32->30
 de [de] <sam-> <np-close> PRP @N< #33->32
 a [o] <-sam> <artd> DET F S @>N #34->35
 Cleópatra [Cleópatra] <cjt-head> <hum> <*> PROP F
 S @P< #35->33
 e [e] <co-prparg> KC @CO #36->35
 Hatshepsut [Hatshepsut] <cjt> <hum> <*> PROP M/F S
 @P< #37->35
 \$(#38->0
 rainhas [rainha] <Htit> <np-close> N F P @N<PRED #39->37
 de [de] <sam-> <np-close> PRP @N< #40->39
 o [o] <-sam> <artd> DET M S @>N #41->42
 Egito [Egito] <cjt-head> <civ> <*> PROP M S @P< #42->40
 \$) #43->0
 \$, #44->0
 mulheres [mulher] <cjt> <Hbio> <Hfam> <Hattr>
 N F P @P< #45->42
 espartanas [espartano] <nat> <jh> <np-close> ADJ
 F P @N< #46->45
 e [e] <co-postnom> KC @CO #47->42
 mulheres [mulher] <cjt> <Hbio> <Hfam> <Hattr>
 <np-close> N F P @N< #48->42
 romanas [romano] <nat> <jh> <np-long> ADJ
 F P @N< #49->45
 \$, #50->0
 apesar=de [apesar=de] <sam-> PRP @<ADVL #51->30
 essas [esse] <dem> <-sam> DET F P @P< #52->51
 não [não] <setop> ADV @<ADVL #53->30
 terem [ter] <mv> V INF 3P @ICL-<SUBJ #54->23
 uma [um] <arti> DET F S @>N #55->56
 vida [vida] <per> N F S @<ACC #56->54
 política [político] <np-close> ADJ F S @N< #57->56
 ativa [ativo] <np-long> ADJ F S @N< #58->56
 \$. #59->0

 Houve [haver] <fmc> <*> <mv> V PS 1/3S IND VFIN @FS-STA #1->0
 progressos [progresso] <cjt-head> <am> N M P @<ACC #2->1
 e [e] <co-acc> KC @CO #3->2

retrocessos [retrocesso] <cjt> <act> N M P @<ACC #4->2
 ao=longo=de [ao=longo=de] <sam-> <np-close> PRP @N< @<ADVL #5->1
 a [o] <artd> <-sam> DET F S @>N #6->7
 história [história] <per> <domain> <sem-r> N F S @P< #7->5
 em [em] <sam-> PRP @<ADVL #8->1
 esse [esse] <-sam> <dem> DET M S @>N #9->10
 âmbito [âmbito] <Labs> N M S @P< #10->8
 \$. #11->0

No=entanto [no=entanto] <kc> <*> ADV @ADVL> #1->10
 \$, #2->0
 em [em] <sam-> PRP @ADVL> #3->10
 o [o] <-sam> <artd> DET M S @>N #4->5
 Brasil [Brasil] <civ> <*> PROP M S @P< #5->3
 \$, #6->0
 o=que [o=que] <clb> <clb-fs> <rel> SPEC M S @SUBJ> #7->9
 se [se] <obj> PERS M/F 3S ACC @ACC> #8->9
 vê [ver] <vH> <mv> V PR 3S IND VFIN @FS-SUBJ> #9->10
 é [ser] <fmc> <mv> V PS 3S IND VFIN @FS-STA #10->0
 que [que] <clb> <clb-fs> KS @SUB #11->14
 as [o] <artd> DET F P @>N #12->13
 mulheres [mulher] <Hbio> <Hfam> <Hattr> N F P @SUBJ> #13->14
 foram [ser] <mv> V PS/MQP 3P IND VFIN @FS-<SC #14->10
 tardiamente [tardiamente] <cjt-head> <deadj> ADV @<ADVL #15->14
 e [e] <co-advl> KC @CO #16->15
 gradativamente [gradativamente] <cjt> <nh> <deadj> ADV
 @<ADVL #17->15
 conquistando [conquistar] <clb> <vH> <mv> V GER
 @ICL-<SC @ICL-<ADVL #18->14
 seus [seu] <poss 3S> <si> DET M P @>N #19->20
 direitos [direito] <domain> <f-right> <mon> N M P
 @<ACC #20->18
 \$. #21->0

A=partir=de [a=partir=de] <*> PRP @ADVL> #1->9
 1827 [1827] <card> <date> <year> NUM M/F P @P< #2->1
 \$, #3->0
 a [o] <artd> DET F S @>N #4->5
 mulher [mulher] <Hbio> <Hfam> <Hattr> N F S
 @SUBJ> #5->7
 brasileira [brasileiro] <nat> <jh> <np-close>
 ADJ F S @N< #6->5
 passou [passar] <cjt-head> <fmc> <aux> V PS 3S
 IND VFIN @FS-STA #7->0
 a [a] PRP @PRT-AUX< #8->7
 ter [ter] <mv> <uninfl> V INF 3S @ICL-AUX< #9->7
 brandos [brando] <jh> ADJ M P @>N #10->11
 direitos [direito] <domain> <f-right> <mon>

N M P @<ACC #11->9
 \$, #12->0
 como [como] <prd> <adv-rel> <np-close> PRP
 @N< @<ADVL #13->9
 permissão [permissão] <asarg> <act-s> N F S
 @P< #14->13
 de [de] <np-close> PRP @N< #15->14
 estudar [estudar] <vH> <mv> <uninfl> V INF 3S
 @ICL-P< #16->15
 \$, #17->0
 e [e] <clb> KC @CO #18->7
 em [em] PRP @ADVL> #19->21
 1888 [1888] <card> <date> <year> NUM M/F S @P< #20->19
 destaca- [destacar] <cjt> <fmc> <hyfen> <mv> <v-pron> V
 PR 3S IND VFIN @FS-STA #21->7
 se [se] <refl> PERS F 3S ACC @<ACC-PASS #22->21
 a [o] <artd> DET F S @>N #23->24
 pioneira [pioneira] <H> N F S @<SUBJ #24->21
 \$' #25->0
 em [em] PRP @<ADVL #26->21
 questões [questão] <act-s> <ac> N F P @P< #27->26
 diplomáticas [diplomático] <nh> <np-close>
 ADJ F P @N< #28->27
 \$' #29->0
 \$, #30->0
 Princesa=Isabel [Princesa=Isabel] <civ> <*>
 <np-close> PROP F S @APP #31->27
 \$, #32->0
 que [que] <clb> <clb-fs> <rel> SPEC M S @SUBJ> #33->34
 assinou [assinar] <vH> <mv> <np-close> V PS 3S
 IND VFIN @FS-N< #34->31
 a [o] <artd> DET F S @>N #35->36
 Lei=Áurea [Lei=Áurea] <tit> <*> PROP F S @<ACC #36->34
 \$, #37->0
 a=qual [o=qual] <clb> <clb-fs> <rel> SPEC F S @SUBJ> #38->39
 aboliu [abolir] <vH> <mv> <np-close> V PS 3S
 IND VFIN @FS-N< #39->36
 a [o] <artd> DET F S @>N #40->41
 escravidão [escravidão] <state-h> N F S @<ACC #41->39
 \$. #42->0

Apesar=de [apesar=de] <*> PRP @ADVL> #1->14
 ela [ela] <clb> PERS F 3S NOM @SUBJ> #2->3
 ter [ter] <mv> V INF 3S @ICL-P< #3->1
 um [um] <arti> DET M S @>N #4->5
 papel [papel] <mat> <cc-r> <ac> N M S @<ACC #5->3
 fundamental [fundamental] <nh> <np-close>
 ADJ M S @N< #6->5

em [em] <sam-> PRP @<ADVL #7->3
 a [o] <-sam> <artd> DET F S @>N #8->9
 história [história] <per> <domain> <sem-r>
 N F S @P< #9->7
 de [de] <sam-> <np-close> PRP @N< #10->9
 o [o] <-sam> <artd> DET M S @>N #11->12
 Brasil [Brasil] <civ> <*> PROP M S @P< #12->10
 \$, #13->0
 demorou [demorar] <nosubj> <cjt-head> <fmc> <mv>
 V PS 3S IND VFIN @FS-STA #14->0
 a [a] PRP @<ACC #15->14
 outras [outro] <diff> <KOMP> DET F P @>N #16->17
 mulheres [mulher] <Hbio> <Hfam> <Hattr> N F P @P< #17->15
 se [se] <clb> <clb-fs> KS @SUB #18->19
 destacarem [destacar] <mv> V FUT 3P SUBJ VFIN
 @FS-<ADVL #19->14
 e [e] <co-fin> <co-fmc> KC @CO #20->14
 \$, #21->0
 após [após] PRP @ADVL> #22->31
 muitas [muito] <quant> DET F P @>N #23->24
 lutas [luta] <cjt-head> <fight> N F P @P< #24->22
 \$, #25->0
 Bertha=Luz [Bertha=Luz] <cjt> <hum> <*> <heur>
 <foreign> PROP F S @P< #26->24
 \$, #27->0
 em [em] PRP @<ADVL #28->19
 1922 [1922] <card> <date> <year> NUM M/F S @P< #29->28
 \$, #30->0
 cria [criar] <nosubj> <cjt> <fmc> <mv> V PR 3S
 IND VFIN @FS-STA #31->14
 a [o] <artd> DET F S @>N #32->33
 Federação=Brasileira [Federação=Brasileira]
 <org> <*> PROP F S @<ACC #33->31
 por [por] <sam-> PRP @<ADVL #34->31
 o [o] <-sam> <artd> DET M S @>N #35->36
 Progresso=Feminino [Progresso=Feminino] <inst> <*>
 <heur> PROP M S @P< #36->34
 \$, #37->0
 que [que] <clb> <clb-fs> KS @SUB #38->39
 foi [ser] <vK> <mv> V PS 3S IND VFIN @FS-<ADVL #39->31
 o [o] <artd> DET M S @>N #40->41
 estopim [estopim] <cord> N M S @<SC #41->39
 para [para] <np-close> PRP @N< #42->41
 a [o] <artd> DET F S @>N #43->44
 conquista [conquista] <act> N F S @P< #44->42
 de [de] <sam-> <np-close> PRP @N< #45->44
 o [o] <-sam> <artd> DET M S @>N #46->47
 voto [voto] <act-s> N M S @P< #47->45

feminino [feminino] <jh> <np-close> ADJ M S @N< #48->47
\$. #49->0

A=partir=de [a=partir=de] <sam-> <*> PRP @ADVL> #1->8
aí [aí] <-sam> <kc> ADV @P< #2->1
\$, #3->0
a [o] <artd> DET F S @>N #4->5
mulher [mulher] <Hbio> <Hfam> <Hattr> N F S
@SUBJ> #5->6
começou [começar] <fmc> <vH> <aux> V PS 3S IND
VFIN @FS-STA #6->0
a [a] PRP @PRT-AUX< #7->6
progredir [progredir] <mv> <uninfl> V INF 3S
@ICL-AUX< #8->6
um=pouco [um=pouco] <quant> ADV @>A #9->10
mais [mais] <quant> <KOMP> ADV @>A #10->11
rapidamente [rapidamente] <deadj> ADV @<ADVL #11->8
\$, #12->0
tendo [ter] <clb> <mv> V GER @ICL-<ADVL #13->8
em [em] PRP @<ADVL #14->13
1935 [1935] <card> <date> <year> NUM M/F S @P< #15->14
a [o] <artd> DET F S @>N #16->19
primeira [primeiro] <NUM-ord> ADJ F S @>N #17->19
parlamentar [parlamentar] <nh> ADJ F S @>N #18->19
eleita [eleger] <n> <vH> V PCP F S @<ACC #19->13
\$, #20->0
Carlota=de=Queiroz [Carlota=de=Queiroz] <hum> <*>
<foreign> <np-close> PROP F S @APP #21->19
\$. #22->0

Conforme [conforme] <com> <*> <adv-rel> PRP
@ADVL> #1->6
a [o] <artd> DET F S @>N #2->3
constituição [constituição] <asarg> <act> N F S
@P< #3->1
\$, #4->0
todos [todo] <quant> DET M P @SUBJ> #5->6
são [ser] <cjt-head> <fmc> <vK> <mv> V PR 3P
IND VFIN @FS-STA #6->0
iguais [igual] <jh> ADJ M P @<SC #7->6
perante [perante] PRP @<ADVL #8->6
a [o] <artd> DET F S @>N #9->10
lei [lei] <conv> <percep-f> N F S @P< #10->8
não [não] <setop> ADV @ADVL> #11->12
havendo [haver] <clb> <mv> V GER @ICL-<ADVL #12->20
mais [muito] <quant> <KOMP> DET F S @>N #13->14
distinção [distinção] <act-d> <am> N F S @<ACC #14->12
de [de] <np-close> PRP @N< #15->14

sexo [sexo] <activity> <f-h> N M S @P< #16->15
 \$, #17->0
 no=entanto [no=entanto] <kc> ADV @ADVL> #18->20
 ainda [ainda] <setop> <atemp> ADV @ADVL> #19->20
 há [haver] <nosubj> <cjt> <fmc> <mv> V PR 3S IND
 VFIN @FS-STA #20->6
 em [em] PRP @<ADVL #21->20
 pleno [pleno] <nh> ADJ M S @>N #22->23
 século [século] <cjt-head> <dur> <per> N M S @P< #23->21
 XXI [XXI] <*> <NUM-ord> <jh> ADJ M S @>N #24->25
 preconceitos [preconceito] <cjt> <asarg> <sem-c>
 N M P @P< #25->23
 machistas [machista] <jh> <np-close> ADJ M P @N< #26->25
 vindos [vir] <mv> <np-close> V PCP M P @ICL-N< #27->25
 tanto [tanto] <parkc-1> ADV @CO #28->31
 de [de] PRP @<ADVL #29->27
 homem [homem] <Hbio> N M S @P< #30->29
 como [como] <parkc-2> ADV @CO #31->30
 de [de] PRP @<SA #32->27
 mulheres [mulher] <Hbio> <Hfam> <Hattr>
 N F P @P< #33->32
 \$. #34->0

Mas [mas] <*> KC @CO #1->4
 isso [isso] <dem> SPEC M S @SUBJ> #2->4
 não [não] <setop> ADV @ADVL> #3->4
 é [ser] <fmc> <vK> <mv> V PR 3S IND VFIN
 @FS-STA #4->0
 problema [problema] <ac> N M S @<SC #5->4
 para [para] PRP @<ADVL #6->4
 abalar [abalar] <mv> <uninfl> V INF 3S
 @ICL-P< #7->6
 a [o] <artd> DET F S @>N #8->10
 boa [bom] ADJ F S @>N #9->10
 realidade [realidade] <sit> <percep-f> N F S
 @<ACC #10->7
 vivida [viver] <vH> <mv> <np-close> V PCP F S
 @ICL-N< #11->10
 por [por] <sam-> PRP @<PASS #12->11
 as [o] <-sam> <artd> DET F P @>N #13->14
 mulheres [mulher] <Hbio> <Hfam> <Hattr> N F P
 @P< #14->12
 \$, #15->0
 já=que [já=que] <clb> <clb-fs> KS @SUB #16->18
 elas [elas] PERS F 3P NOM @SUBJ> #17->18
 são [ser] <cjt-head> <vK> <mv> V PR 3P IND VFIN
 @FS-<ADVL #18->11
 maioria [maioria] <HH> N F S @<SC #19->18

em [em] <sam-> <np-close> PRP @N< #20->19
 o [o] <-sam> <artd> DET M S @>N #21->22
 país [país] <Lciv> N M S @P< #22->20
 e [e] <co-fin> KC @CO #23->18
 em [em] <sam-> PRP @<ADVL #24->18
 o [o] <-sam> <artd> DET M S @>N #25->26
 eleitorado [eleitorado] <HH> N M S @P< #26->24
 uma=vez=que [uma=vez=que] <clb> <clb-fs> KS @SUB #27->32
 70 [70] <card> NUM M P @>N #28->29
 milhões [milhão] <amount> N M P @SUBJ> #29->32
 de [de] <np-close> PRP @N< #30->29
 mulheres [mulher] <Hbio> <Hfam> <Hattr> N F P @P< #31->30
 votam [votar] <cjt> <vH> <mv> V PR 3P IND VFIN @FS-<ADVL #32->18
 sendo [ser] <clb> <vK> <mv> V GER @ICL-<ADVL #33->32
 cinco=milhões [cinco=milhões] <cjt-head> <card>
 NUM M P @<SC #34->33
 a [o] <artd> DET F S @>N #35->36
 mais [muito] <quant> <KOMP> DET M/F S/P @<SUBJ #36->33
 que [que] KS @SUB #37->0
 os [o] <artd> DET M P @>N #38->39
 homens [homem] <Hbio> N M P @PRED> #39->0
 \$, #40->0
 segundo [segundo] <cjt> <NUM-ord> <jh> ADJ M S @<SC #41->34
 a [o] <artd> DET F S @>N #42->43
 Agência=Nacional [Agência=Nacional] <cjt> <org> <*>
 PROP F S @<SC #43->34
 \$. #44->0

O=sexo=frágil [o=sexo=frágil] <HH> <*> N M S @SUBJ> #1->6
 com [com] <np-close> PRP @N< #2->1
 certeza [certeza] <am> N F S @P< #3->2
 já [já] <setop> <atemp> ADV @ADVL> #4->6
 se [se] <refl> PERS M/F 3S ACC @ACC> #5->6
 tornou [tornar] <cjt-head> <fmc> <mv> <v-pron> V PS 3S
 IND VFIN @FS-STA #6->0
 forte [forte] <jh> ADJ M/F S @<OC #7->6
 \$, #8->0
 não [não] <setop> ADV @>A #9->10
 só [só] <setop> ADV @>P #10->11
 em [em] <sam-> PRP @<ADVL #11->6
 a [o] <-sam> <artd> DET F S @>N #12->13
 política [política] <domain> <sem-c> N F S @P< #13->11
 \$, #14->0
 mas [mas] <clb> KC @CO #15->6
 em [em] <sam-> PRP @ADVL> #16->22
 as [o] <-sam> <artd> DET F P @>N #17->18
 universidades [universidade] <inst> N F P @P< #18->16
 as [o] <artd> DET F P @>N #19->20

mulheres [mulher] <Hbio> <Hfam> <Hattr> N F P @SUBJ> #20->21
 estão [estar] <cjt> <fmc> <aux> V PR 3P IND VFIN @FS-STA #21->6
 ascendendo [ascender] <clb> <mv> V GER @ICL-AUX< #22->21
 cada=vez=mais [cada=vez=mais] <quant> ADV @<ADVL #23->22
 em=relação=a [em=relação=a] <sam-> PRP @A< @<ADVL #24->22
 os [o] <artd> <-sam> DET M P @>N #25->26
 homens [homem] <Hbio> N M P @P< #26->24
 \$, #27->0
 e [e] <clb> KC @CO #28->6
 em [em] <sam-> PRP @ADVL> #29->37
 as [o] <-sam> <artd> DET F P @>N #30->32
 últimas [último] <NUM-ord> ADJ F P @>N #31->32
 décadas [década] <dur> N F P @P< #32->29
 as [o] <artd> DET F P @>N #33->34
 mulheres [mulher] <Hbio> <Hfam> <Hattr> N F P
 @<SUBJ #34->21
 está [estar] <cjt> <fmc> <aux> V PR 3S IND VFIN
 @FS-STA #35->6
 se [se] <refl> PERS M 3S/P ACC @ACC> #36->37
 matriculando [matricular] <clb> <vH> <mv> <v-pron>
 V GER @ICL-AUX< #37->35
 22 [22] <card> NUM M P @>N #38->39
 \$% [%] N M P @<SUBJ #39->35
 a [o] <artd> DET F S @>N #40->41
 mais [muito] <quant> <KOMP> DET M/F S/P @<SC #41->37
 que [que] KS @SUB #42->0
 os [o] <artd> DET M P @>N #43->44
 homens [homem] <Hbio> N M P @PRED> #44->0
 \$, #45->0
 segundo [segundo] <com> <adv-rel> <np-close>
 PRP @N< #46->41
 Instituto=Nacional=de=Estudos [Instituto=Nacional=de=Estudos]
 <cjt-head> <tit> <asarg> <*> PROP M S @P< #47->46
 e [e] <co-prparg> KC @CO #48->47
 Pesquisas=Educacionais [Pesquisas=Educacionais] <cjt> <hum>
 <asarg> <*> <heur> PROP M/F S @P< #49->47
 \$(#50->0
 Inep [Inep] <top> <*> <heur> <foreign> <np-close>
 PROP M/F S @N<PRED #51->41
 \$) #52->0
 \$. #53->0

 No=entanto [no=entanto] <kc> <*> ADV @ADVL> #1->9
 em [em] <sam-> PRP @ADVL> #2->9
 essa [esse] <-sam> <dem> DET F S @>N #3->4
 maré [maré] <event> N F S @P< #4->2
 de=rosas [de=rosas] <pp> ADJ M/F S/P @N< #5->4
 \$, #6->0

as [o] <artd> DET F P @>N #7->8
 mulheres [mulher] <Hbio> <Hfam> <Hattr> N F P
 @SUBJ> #8->9
 sofrem [sofrer] <nosubj> <cjt-head> <fmc> <vH> <mv>
 V PR 3P IND VFIN @FS-STA #9->0
 com [com] PRP @<PIV #10->9
 a [o] <artd> DET F S @>N #11->12
 discriminação [discriminação] <activity> <am> N F S
 @P< #12->10
 quanto=a [quanto=a] <sam-> <np-close> PRP @N<
 @<ADVL #13->9
 o [o] <-sam> <artd> DET M S @>N #14->15
 salário [salário] <mon> N M S @P< #15->13
 \$, #16->0
 pois [pois] <clb> KS @ADVL> #17->19
 muitas [muito] <quant> DET F P @SUBJ> #18->19
 exercem [exercer] <cjt> <fmc> <vH> <mv> V PR 3P
 IND VFIN @FS-STA #19->9
 cargos [cargo] <ac> N M P @<ACC #20->19
 maiores [grande] <KOMP> <np-close> ADJ
 M P @N< #21->20
 que [que] <clb> KS @SUB #22->26
 de [de] PRP @ADVL> #23->26
 homens [homem] <Hbio> N M P @P< #24->23
 e [e] <co-fin> <co-fmc> KC @CO #25->9
 ganham [ganhar] <nosubj> <cjt> <fmc> <vH> <mv> <nosubj-safe>
 V PR 3P IND VFIN @FS-STA #26->9
 salários [salário] <mon> N M P @<ACC #27->26
 inferiores [inferior] <KOMP> <SUP> <np-close> ADJ M
 P @N< #28->27
 \$. #29->0

Mas [mas] <*> KC @CO #1->5
 paulatinamente [paulatinamente] <nh> <deadj> ADV
 @ADVL> #2->6
 essas [esse] <dem> DET F P @>N #3->4
 guerreiras [guerreira] <Hprof> N F P @SUBJ> #4->5
 vão [ir] <fmc> <aux> V PR 3P IND VFIN @FS-STA #5->0
 enfrentar [enfrentar] <cjt-head> <vH> <mv> <uninfl> V
 INF 3P @ICL-AUX< #6->5
 e [e] KC @CO #7->9
 \$, #8->0
 com [com] PRP @<PRED @ADVL> #9->6
 certeza [certeza] <am> N F S @P< #10->9
 \$, #11->0
 vencer [vencer] <cjt> <vH> <mv> <uninfl> V INF 3P
 @ICL-AUX< #12->6
 mais [mais] <quant> <KOMP> ADV @<ADVL #13->12

essa [esse] <dem> DET F S @>N #14->15
batalha [batalha] <occ> N F S @<ACC #15->12
\$! #16->0

APÊNDICE C – TIPOS DE CORREFERÊNCIAS DO CORP

Tabela C.1: Lista de regras do CORP.

Código	Descrição
Casamento de Padrões Exato	[o Brasil], [o Brasil]
Casamento Parcial pelo Núcleo	[o ministro da justiça], [o ministro]
Aposto Explicativo	[A ministra da justiça do país], [Elisabete Guigou], ...
Aposto Especificativo	[o arqueólogo português], [Francisco Alves]
Acrônimo	[Organização das Nações Unidas], [a ONU]
Predicado Nominativo	[A França] é [o único país que se recusa a aceitar] ...
Pronome Relativo	[Wilkinson Microwave Anisotropy Probe], [cujos] primeiros ...
Casamento Restrito pelo Núcleo1	[Universidade Federal de São Paulo] ... [a Universidade] ...
Casamento Restrito pelo Núcleo2	[O menino de pijama listrado], [o pijama listrado].
Casamento entre Nomes Próprios	[Califórnia],[a região sul da Califórnia].
Casamento Parcial entre Nomes Próprios	[o agrônomo da UFSC, Miguel Guerra] e [Guerra]
Hiponímia	... como [as abelhas] fabricam mel? [Os insetos] saem ...
Sinonímia	[A Terra] é [um astro do sistema solar]. [Esse planeta] ...

Fonte: Adaptado de (Fonseca et al., 2018b)