

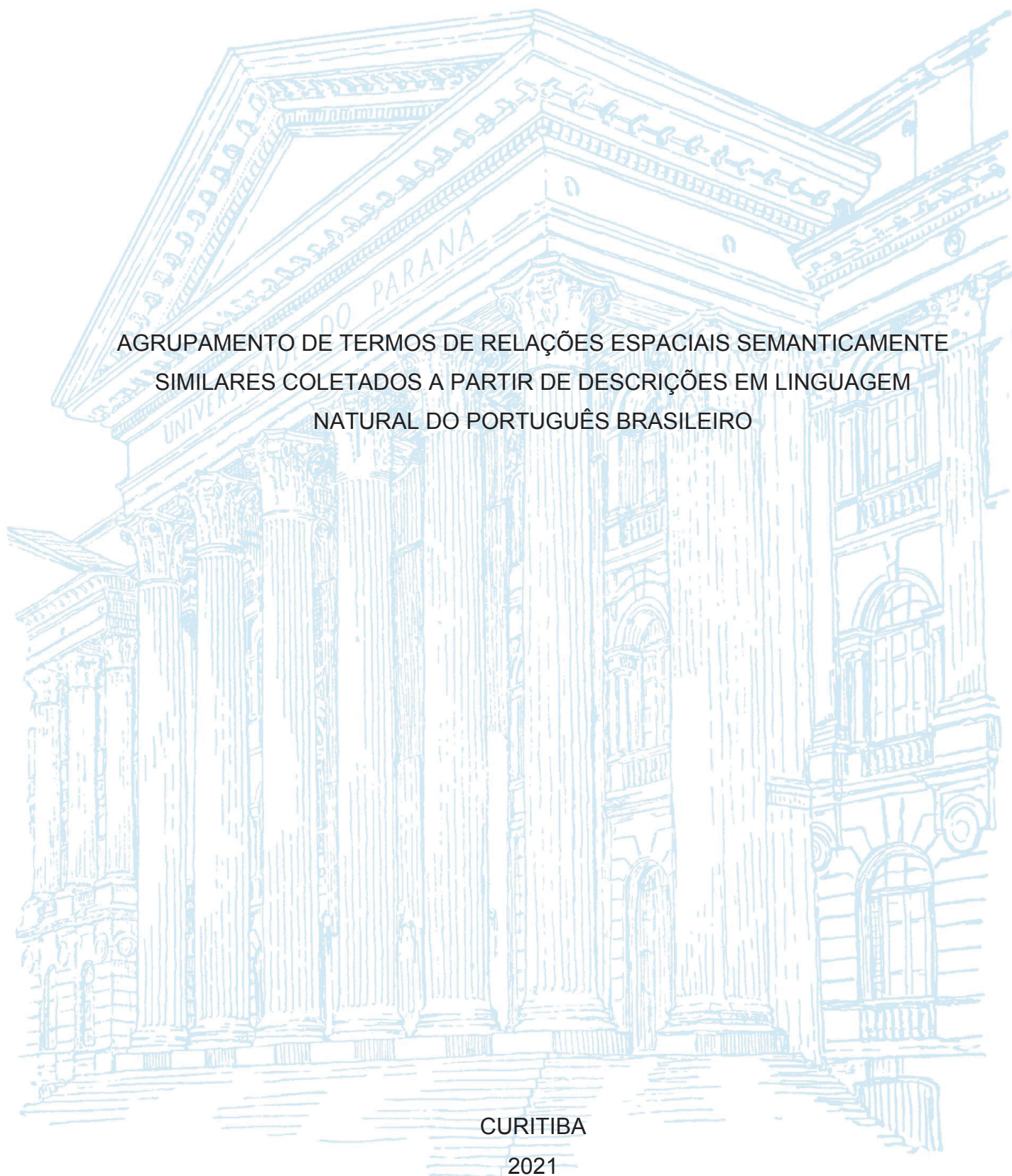
UNIVERSIDADE FEDERAL DO PARANÁ

CRISTIANE KUTIANSKI MARCHI FAGUNDES

AGRUPAMENTO DE TERMOS DE RELAÇÕES ESPACIAIS SEMANTICAMENTE
SIMILARES COLETADOS A PARTIR DE DESCRIÇÕES EM LINGUAGEM
NATURAL DO PORTUGUÊS BRASILEIRO

CURITIBA

2021



CRISTIANE KUTIANSKI MARCHI FAGUNDES

AGRUPAMENTO DE TERMOS DE RELAÇÕES ESPACIAIS SEMANTICAMENTE
SIMILARES COLETADOS A PARTIR DE DESCRIÇÕES EM LINGUAGEM
NATURAL DO PORTUGUÊS BRASILEIRO

Tese apresentada ao Programa de Pós-Graduação em Ciências Geodésicas, Departamento de Geomática, área de concentração em Cartografia, Setor de Ciências da Terra, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Doutor em Ciências Geodésicas.

Orientadora: Profa. Dra. Luciene Stamato Delazari

CURITIBA

2021

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

F156a Fagundes, Cristiane Kutianski Marchi
Agrupamento de termos de relações espaciais semanticamente similares coletados a partir de descrições em linguagem natural do português brasileiro [recurso eletrônico] / Cristiane Kutianski Marchi Fagundes. – Curitiba, 2021.

Tese - Universidade Federal do Paraná, Setor de Ciências da Terra, Programa de Pós-Graduação em Ciências Geodésicas, 2021.

Orientadora: Luciene Stamato Delazari.

1. Processamento de linguagem natural (Computação). 2. Anúncios ao ar livre. 3. Sistemas de informação geográfica. I. Universidade Federal do Paraná. II. Delazari, Luciene Stamato. III. Título.

CDD: 910.285

Bibliotecária: Vanusa Maciel CRB- 9/1928



MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIÊNCIAS DA TERRA
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO CIÊNCIAS
GEODÉSICAS - 40001016002P6

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em CIÊNCIAS GEODÉSICAS da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **CRISTIANE KUTIANSKI MARCHI FAGUNDES** intitulada: **AGRUPAMENTO DE TERMOS DE RELAÇÕES ESPACIAIS SEMANTICAMENTE SIMILARES COLETADOS A PARTIR DE DESCRIÇÕES EM LINGUAGEM NATURAL DO PORTUGUÊS BRASILEIRO**, sob orientação da Profa. Dra. LUCIENE STAMATO DELAZARI, que após terem inquirido a aluna e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutor 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, 10 de Maio de 2021.

Assinatura Eletrônica

11/05/2021 13:02:33.0

LUCIENE STAMATO DELAZARI

Presidente da Banca Examinadora

Assinatura Eletrônica

11/05/2021 13:30:19.0

JULIA CELIA MERCEDES STRAUCH

Avaliador Externo (ESCOLA NACIONAL DE CIÊNCIAS ESTATÍSTICAS
)

Assinatura Eletrônica

11/05/2021 14:01:47.0

MARCIO AUGUSTO REOLON SCHMIDT

Avaliador Interno (UNIVERSIDADE FEDERAL DE UBERLÂNDIA)

Assinatura Eletrônica

11/05/2021 15:24:33.0

ALFREDO PEREIRA DE QUEIROZ FILHO

Avaliador Externo (FFLCH DA UNIVERSIDADE DE SÃO PAULO)

Centro Politécnico - Caixa Postal 19001 - CURITIBA - Paraná - Brasil

CEP 81531-980 - Tel: (41) 3361-3153 - E-mail: cpgcg@ufpr.br

Documento assinado eletronicamente de acordo com o disposto na legislação federal Decreto 8539 de 08 de outubro de 2015.

Gerado e autenticado pelo SIGA-UFPR, com a seguinte identificação única: 91095

Para autenticar este documento/assinatura, acesse <https://www.prppg.ufpr.br/siga/visitante/autenticacaoassinaturas.jsp>
e insira o código 91095

Dedico este trabalho ao meu marido Persio, por toda a compreensão, apoio, incentivo e amor que recebi durante toda a caminhada do Doutorado. Quatro palavras importantes que recebem e receberão minha eterna gratidão.

AGRADECIMENTOS

Aos meus pais Maria e Nelino, primeiramente pela minha vida e também pelo esforço dispendido na minha criação e educação.

Às minhas filhas Amanda e Bárbara pela paciência, entendimento, incentivo e principalmente pelo amor envolvido nestes anos dispensados a este trabalho. Ao meu pequeno Cassiano (*in memoriam*), que mesmo não estando aqui, tenho certeza que me acompanha e me ilumina em todos os lugares e situações da minha vida.

À minha irmã Luciane pelo “*help in English*” e pelas palavras de incentivo. À minha madrinha Isabel que sempre torce pelas minhas conquistas.

À minha orientadora Professora Dr^a. Luciene Stamato Delazari, pela confiança, apoio, conselhos e principalmente pela orientação.

Aos professores do Programa de Pós-Graduação em Ciências Geodésicas, pela dedicação e qualidade das aulas ministradas.

Às minhas amigas Maria Helena, Edna, Joelma, Lúcia, Regiane, Evandra, Adriana e Preeti que me acompanharam durante o desenvolvimento da pesquisa e me deram apoio para continuar até o final.

À Professora Dr^a. Kristin Stock que fez a orientação dos meus trabalhos na Massey University em Auckland, Nova Zelândia.

Aos funcionários, em especial à secretaria da Pós-Graduação em Ciências Geodésicas pelo pronto atendimento e colaboração.

Aos professores e integrantes da banca examinadora pelo comparecimento, colaboração e sugestões para o desenvolvimento da pesquisa.

À Universidade Federal do Paraná e ao Programa de Pós-Graduação em Ciências Geodésicas pelo aceite da pesquisa e pela infraestrutura fornecida.

À CAPES pela bolsa de pesquisa pelo período de 05/2016 a 02/2018 e pelo fomento do Programa de Doutorado Sanduíche no Exterior – 88881.189171/2018-01, pelo período de 12/2018 a 05/2019 na Massey University, Nova Zelândia.

Ao CNPQ pela bolsa de pesquisa processo 141254/2018-1 pelo período de 03/2018 a 11/2018 e 06/2019 a 04/2020.

A todos que contribuíram direta ou indiretamente com o desenvolvimento desta pesquisa, sendo desde pensamentos positivos até palavras motivadoras; e principalmente, que participaram ou acompanharam alguns momentos desta caminhada chamada “doutorado”.

“Se, a princípio, a ideia não é absurda, então não há esperanças para ela.”

Albert Einstein

RESUMO

Estudos em linguagem natural abordam a extração de pontos de referência e feições geográficas que são úteis em banco de dados para consultas de lugares, parques, praças, entre outros. Porém, estudos com termos de relações espaciais que fazem parte das descrições de localização do cotidiano das pessoas ainda são voltados para determinadas preposições e são explorados principalmente na língua inglesa. Os termos de relações espaciais topológicos também ganharam espaço nos estudos nas últimas décadas em função da robótica e de aplicações em Sistemas de Informações Geográficas. No entanto, necessita-se de mais estudos exploratórios em outras línguas e com outros tipos de relações espaciais, que não somente os topológicos. Desta maneira, poderá haver melhorias nas aplicações computacionais voltadas para a linguagem natural como similaridades semânticas e ontologias e em aplicações geoespaciais. Para tanto, esta pesquisa partiu da hipótese principal de que é possível agrupar os termos de relações espaciais descritas em linguagem natural por meio de um conjunto reduzido de termos semanticamente similares a partir de ambientes reais do cotidiano das pessoas. Para isto, foi desenvolvida uma metodologia para criação de grupos semanticamente similares de termos de relações espaciais utilizados na linguagem natural do português brasileiro. Para tanto foram realizados três testes. Dois testes para coleta, análise e verificação dos termos mais utilizados, sendo um no contexto *outdoor* e outro no contexto *indoor* com a proposição de duas hipóteses: 1) pelo grau de liberdade que a linguagem natural oferece, os termos de relações espaciais que se destacam são diferentes para cada tipo de ambiente; 2) se as pessoas fazem descrições de localização em linguagem natural em um determinado contexto e ambiente, de acordo com suas percepções do ambiente e do espaço e são livres para optar pelos elementos a serem relacionados, então elas utilizam os termos de relações espaciais no referencial centrado nos objetos. O terceiro teste de associação livre foi realizado para obtenção dos agrupamentos iniciais de termos de relações espaciais com a proposição de mais duas hipóteses: 1) se os termos semanticamente similares de relações espaciais encontrados através do mesmo *relatum* e do mesmo *locatum* dos dois primeiros testes coincidem com os termos encontrados no teste de associação livre; 2) se é possível criar uma classe de termos semanticamente similares. Para a obtenção dos agrupamentos finais foram elaboradas e aplicadas cinco regras para contribuir e facilitar o processo de identificação e análise dos grupos de relações espaciais utilizados como semanticamente similares na língua portuguesa brasileira. Além do que, espera-se que tais regras possam ser aplicadas para outras línguas. Um dos resultados mostra que foi possível verificar o distanciamento dos termos similares pela quantidade de citações de cada termo em ordem decrescente. Isto facilita a utilização dos termos em aplicações computacionais que necessitem a utilização de mais de um termo, como por exemplo, a geração de descrições de localização em linguagem natural. Além disso, os termos encontrados podem ser utilizados no conjunto de interfaces do projeto *Where am I? (parceria UFPR e Massey University da Nova Zelândia)* respeitando as classificações de cada termo.

Palavras-chave: Linguagem natural. Descrições de localização. Expressões locativas. *Outdoor*. *Indoor*.

ABSTRACT

Studies in natural language address the extraction of landmarks and geographic features that are useful in databases for queries of places, parks, squares, among others. And studies with terms of spatial relations that are part of the descriptions of people's daily lives are still focused on certain prepositions and are explored mainly in the English language. Additionally, terms of topological spatial relations have also gained space in studies in the last decades due to robotics and applications in Geographic Information Systems. However, more exploratory studies are needed in other languages and with other types of spatial relations, not just topological ones. In this way, there may be improvements in computational applications for natural language such as semantic similarities, ontologies and in geospatial applications. To this end, this research started from the main hypothesis that it is possible to group the terms of spatial relations described in natural language through a reduced set of semantically similar terms from real environments of people's daily lives. For this, a methodology was developed to create semantically similar groups of terms of spatial relations used in the natural language of Brazilian Portuguese. For this purpose, three tests were carried out. Two tests were for collection, analysis and verification of the most used terms, one in the outdoor context and the other in the indoor context with the proposition of two hypotheses: 1) by the freedom that natural language offers, the terms of spatial relations that stand out in the descriptions are different for each type of environment; 2) if people make descriptions of location in natural language in a given context and environment, according to their perceptions of the environment and space and they are free to choose the elements to be related, then they use the terms of relations in the object-centered framework. The third was the free association test and it was performed to obtain the initial groups of terms of spatial relations. To this test was made the proposition of two more hypotheses: 1) if the semantically similar terms of spatial relations found through the same *relatum* and the same *locatum* of the first two tests coincide with the terms found in the free association test; 2) it is possible to create a class of semantically similar terms. To obtain the final groups, five rules were elaborated and applied to contribute and facilitate the process of identification and analysis of groups of spatial relations used as semantically similar in the Brazilian Portuguese language. In addition, it is hoped that such rules can be applied to other languages. One of the results shows that it was possible to verify the distance from similar terms by accounting for the number of citations for each term in decreasing order. This result facilitates the use of terms in computational applications that require the use of more than one term, such as, for example, the generation of location descriptions in natural language. In addition, the terms found can be used in the set of interfaces of the Where am I? project (the agreement of UFPR and Massey University of New Zealand) respecting the classification of each term.

Keywords: Natural language. Location descriptions. Locative expression. *Outdoor*. *Indoor*.

LISTA DE FIGURAS

FIGURA 1 - RELAÇÕES ESPACIAIS PARA UMA SITUAÇÃO GEOGRÁFICA.....	41
FIGURA 2 - MODELOS PARA AS RELAÇÕES ESPACIAIS “SOBREPOSIÇÃO” E “ESTAR DENTRO”	50
FIGURA 3 - COGNIÇÃO SOBRE DIFERENTES CORPOS D'ÁGUA	56
FIGURA 4 – TRIÂNGULO DO SIGNIFICADO	61
FIGURA 5 – CONCEITOS DE VIZINHANÇA PARA RELAÇÕES TOPOLÓGICAS..	68
FIGURA 6 – RELAÇÃO DE DISTÂNCIA PARA RELAÇÕES TOPOLÓGICAS.....	68
FIGURA 7 – RELAÇÃO DE DIREÇÃO PARA RELAÇÕES TOPOLÓGICAS	69
FIGURA 8 - ESQUEMA DA METODOLOGIA	81
FIGURA 9 - FOTOGRAFIAS DA ÁREA CENTRAL DE CURITIBA (BRA1)	84
FIGURA 10 - FOTOGRAFIAS DA ÁREA RESIDENCIAL DE UBERLÂNDIA (BRA4)	84
FIGURA 11 - FOTOGRAFIAS DA ÁREA CENTRAL DA NOVA ZELÂNDIA (NZ3)...	85
FIGURA 12 - FOTOGRAFIAS DA ÁREA NÃO URBANIZADA DA NOVA ZELÂNDIA (NZ1).....	85
FIGURA 13 - LOCALIZAÇÃO DAS ÁREAS <i>INDOOR</i>	87
FIGURA 14 - POSIÇÃO A (EDIFÍCIO I) ANDAR TÉRREO DO BLOCO DE ENGENHARIA MECÂNICA.....	88
FIGURA 15 - POSIÇÃO B (EDIFÍCIO II) SEGUNDO ANDAR DO BLOCO DA ADMINISTRAÇÃO	89
FIGURA 16 - FLUXOGRAMA DAS REGRAS	94
FIGURA 17 - AGRUPAMENTO DOS TERMOS DE RELAÇÕES ESPACIAIS	95
FIGURA 18 – VISUALIZAÇÃO ESPACIAL DOS <i>LOCATA</i> – BRA1	101
FIGURA 19 – VISUALIZAÇÃO ESPACIAL DOS <i>RELATA</i> – BRA1	102
FIGURA 20 - ESPACIALIZAÇÃO DOS TERMOS DE RELAÇÕES ESPACIAIS – BRA1.....	113
FIGURA 21 - ESPACIALIZAÇÃO DOS TERMOS DE RELAÇÕES ESPACIAIS – BRA2.....	114
FIGURA 22 - PONTO DE VISTA - LOCALIDADE BRA3	136
FIGURA 23 - PONTO DE VISTA - LOCALIDADE NZ2	137
FIGURA 24 – ESCOLARIDADE DOS PARTICIPANTES.....	139
FIGURA 25 – FAIXA ETÁRIA DOS PARTICIPANTES	139

FIGURA 26 - RELACIONAMENTOS DO AGRUPAMENTO INICIAL ENTRE OS TERMOS DO GRUPO "EM FRENTE"	141
FIGURA 27 - RESULTADO DA APLICAÇÃO DE TODAS AS REGRAS PARA OS TERMOS "EM FRENTE" E "DE FRENTE"	158
FIGURA 28 - DISTÂNCIA ENTRE OS TERMOS SIMILARES DO AGRUPAMENTO "EM FRENTE" E "DE FRENTE"	162
FIGURA 29 - REDE DE TERMOS DE RELAÇÕES ESPACIAIS	167

LISTA DE GRÁFICOS

GRÁFICO 1 - CLASSIFICAÇÃO DOS TERMOS DE RELAÇÕES ESPACIAIS TESTE <i>ONLINE</i>	106
GRÁFICO 2 - CLASSIFICAÇÃO DOS TERMOS DE RELAÇÕES ESPACIAIS – <i>INDOOR</i>	120
GRÁFICO 3 - CLASSIFICAÇÕES ÚNICAS – LOCALIDADES DO BRASIL	128
GRÁFICO 4 - CLASSIFICAÇÕES ÚNICAS - LOCALIDADES DA NOVA ZELÂNDIA	128
GRÁFICO 5 - CLASSIFICAÇÕES CONJUNTAS - LOCALIDADES DO BRASIL....	129
GRÁFICO 6 - CLASSIFICAÇÕES CONJUNTAS - LOCALIDADES DA NOVA ZELÂNDIA.....	130
GRÁFICO 7 - CLASSIFICAÇÃO DOS TERMOS DE RELAÇÕES ESPACIAIS DOS MESMOS <i>LOCATA</i> E <i>RELATA</i> – <i>INDOOR</i>	131
GRÁFICO 8 - RELAÇÕES ESPACIAIS MAIS UTILIZADAS NAS ÁREAS DO BRASIL	132
GRÁFICO 9 - RELAÇÕES ESPACIAIS MAIS UTILIZADAS NAS ÁREAS DA NOVA ZELÂNDIA.....	132
GRÁFICO 10 - RELAÇÕES ESPACIAIS MAIS UTILIZADAS NO CONTEXTO <i>OUTDOOR</i>	133
GRÁFICO 11 - RELAÇÕES ESPACIAIS MAIS UTILIZADAS NAS DUAS POSIÇÕES DO <i>INDOOR</i>	134
GRÁFICO 12 - RELAÇÕES ESPACIAIS MAIS UTILIZADAS NO CONTEXTO <i>INDOOR</i>	135
GRÁFICO 13 - RESULTADO DA PRIMEIRA REGRA APLICADA AO TERMO “EM FRENTE”	143
GRÁFICO 14 - RESULTADO DA SEGUNDA REGRA APLICADA AO GRUPO “EM FRENTE À”	144

LISTA DE QUADROS

QUADRO 1 - EXEMPLO DA ESTRUTURA DO TESTE DE ASSOCIAÇÃO DE PALAVRAS	92
QUADRO 2 - SUBSTITUIÇÕES PELO TERMO "À FRETE"	110
QUADRO 3 - SUBSTITUIÇÃO PELO TERMO "EM DIREÇÃO"	110
QUADRO 4 - SUBSTITUIÇÃO PELO TERMO "COM"	111
QUADRO 5 - SUBSTITUIÇÃO PELO TERMO "À FRENTE" - <i>INDOOR</i>	121
QUADRO 6 - SUBSTITUIÇÃO PELO TERMO "EM DIREÇÃO" - <i>INDOOR</i>	121
QUADRO 7 - SUBSTITUIÇÃO PELO TERMO "DE ... PARA..."	121
QUADRO 8 - SUBSTITUIÇÃO PELOS TERMOS "TEM/EXISTE/HÁ"	122
QUADRO 9 - SUBSTITUIÇÃO PELOS TERMOS "NO/NA"	122
QUADRO 10 - SUBSTITUIÇÃO PELO TERMO "À FRENTE"	122

LISTA DE TABELAS

TABELA 1 - TOTAL DE TERMOS DE RELAÇÕES ESPACIAIS CONTEXTO <i>OUTDOOR</i>	97
TABELA 2 - DESMEMBRAMENTO RELAÇÕES ESPACIAIS - LOCALIDADE BRA1	100
TABELA 3 - AGRUPAMENTO DO TERMO DE RELAÇÃO ESPACIAL “EM FRENTE” DO TESTE <i>ONLINE</i>	104
TABELA 4 - AGRUPAMENTO E CLASSIFICAÇÃO DOS TERMOS DE RELAÇÕES ESPACIAIS DO TESTE <i>ONLINE</i>	105
TABELA 5 - AGRUPAMENTO DOS VERBOS DO TESTE <i>ONLINE</i>	107
TABELA 6 - RELAÇÕES ESPACIAIS UTILIZANDO O TERMO “COM”	109
TABELA 7 - RELAÇÕES ESPACIAIS A PARTIR DOS MESMOS <i>LOCATA</i> E <i>RELATA</i> – TESTE <i>ONLINE</i> - BRA1.....	112
TABELA 8 - DESMEMBRAMENTO RELAÇÕES ESPACIAIS – POSIÇÃO A.....	117
TABELA 9 - AGRUPAMENTO DO TERMO DE RELAÇÃO ESPACIAL “EM FRENTE” DO TESTE PRESENCIAL.....	118
TABELA 10 - AGRUPAMENTO E CLASSIFICAÇÃO DOS TERMOS DE RELAÇÕES ESPACIAIS DO TESTE PRESENCIAL.....	119
TABELA 11 - AGRUPAMENTO DOS VERBOS DO TESTE PRESENCIAL	120
TABELA 12 - RELAÇÕES ESPACIAIS A PARTIR DOS MESMOS <i>RELATA</i> E <i>LOCATA</i> – TESTE PRESENCIAL - POSIÇÃO A.....	123
TABELA 13 – NÚMEROS GERAIS DA COLETA DE DADOS	124
TABELA 14 - TERMOS DE RELAÇÕES ESPACIAIS <i>OUTDOOR</i> E <i>INDOOR</i>	125
TABELA 15 – EIXO DE REFERÊNCIA - <i>OUTDOOR</i>	137
TABELA 16 - EGOCENTRISMO TESTE <i>INDOOR</i>	138
TABELA 17 - AGRUPAMENTO INICIAL DOS TERMOS DE RELAÇÕES ESPACIAIS - “EM FRENTE”	140
TABELA 18 – RESULTADO DAS DUAS PRIMEIRAS REGRAS.....	144
TABELA 19 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “EM FRENTE À E DE FRENTE AO”	147
TABELA 20 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “EM FRENTE À E À FRENTE DE”	147

TABELA 21 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “NA E EM UM”.....	148
TABELA 22 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “NA E ENTRE”	148
TABELA 23 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “PERTO DO” E “PRÓXIMO DE”	148
TABELA 24 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “COM” E “QUE TEM”	149
TABELA 25 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “DE FRENTE AO E À FRENTE DE”	149
TABELA 26 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “DO OUTRO LADO E À FRENTE DE”	150
TABELA 27 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “DO OUTRO LADO E AO LADO DE”	150
TABELA 28 - RESULTADO DA QUARTA REGRA APLICADA AOS GRUPOS “EM FRENTE À E DE FRENTE AO”	151
TABELA 29 - RESULTADO DA QUARTA REGRA APLICADA AOS GRUPOS “EM FRENTE À E À FRENTE DE”	151
TABELA 30 - RESULTADO DA QUARTA REGRA APLICADA À UNIÃO DOS GRUPOS	152
TABELA 31 - RESULTADO DA QUINTA REGRA APLICADA À UNIÃO DOS GRUPOS	153
TABELA 32 - TERMOS SEMANTICAMENTE SIMILARES ANALISADOS POR NATIVO DA LÍNGUA	159
TABELA 33 - TERMOS SEMANTICAMENTE SIMILARES - <i>OUTDOOR</i>	163
TABELA 34 - TERMOS SEMANTICAMENTE SIMILARES - <i>INDOOR</i>	163
TABELA 35 - TERMOS COM CLASSIFICAÇÕES DIFERENTES - <i>OUTDOOR</i> E <i>INDOOR</i>	164
TABELA 36 - TERMOS DIFERENTES DE RELAÇÕES ESPACIAIS PARA OS MESMOS <i>LOCATA</i> E <i>RELATA</i> - <i>OUTDOOR</i>	165
TABELA 37 - TERMOS DIFERENTES DE RELAÇÕES ESPACIAIS PARA OS MESMOS <i>LOCATA</i> E <i>RELATA</i> - <i>INDOOR</i>	165
TABELA 38 - CARACTERÍSTICAS DAS RELAÇÕES ESPACIAIS	169

TABELA 39 – ESTRUTURAS DE EXPRESSÕES LOCATIVAS – TERMO “EM FRENTE” – EIXO EGOCÊNTRICO	171
TABELA 40 – ESTRUTURAS DE EXPRESSÕES LOCATIVAS – TERMO EM FRENTE – EIXO CENTRADO NOS ELEMENTOS	172

LISTA DE ABREVIATURAS OU SIGLAS

PLN	- Processamento de Linguagem Natural
SIG	- Sistemas de Informações Geográficas
IDE	- Infraestrutura de Dados Espaciais
WEB	- <i>World Wide Web</i>
INDE	- Infraestrutura Nacional de Dados Espaciais do Brasil
BDG	- Banco de Dados Geográficos
UFPR	- Universidade Federal do Paraná
OWL	- <i>Ontology Web Language</i>
W3C	- <i>World Wide Web Consortium</i>
RDF	- <i>Resource Description Framework</i>
XML	- <i>Extensible Markup Language</i>
GNSS	- Sistema Global de Navegação por Satélite

SUMÁRIO

1 INTRODUÇÃO	16
1.1 JUSTIFICATIVA	22
1.2 OBJETIVO GERAL	28
1.3 OBJETIVOS ESPECÍFICOS	28
1.4 ESTRUTURA E CONTEÚDO.....	29
2 REVISÃO DE LITERATURA	30
2.1 LINGUAGENS NATURAL E ESPACIAL	30
2.1.1 Inglês, português e outras línguas	35
2.2 RELAÇÕES ESPACIAIS	39
2.2.1 Eixos de referência espacial.....	44
2.2.2 Extração de relações espaciais.....	48
2.2.3 Estudos interlinguísticos de descrições de localização	51
2.3 SEMÂNTICA	53
2.4 SIMILARIDADE SEMÂNTICA	62
2.5 ONTOLOGIA.....	70
2.6 RESUMO DA REVISÃO DE LITERATURA.....	77
3 MATERIAIS E MÉTODOS	80
3.1 MATERIAIS	80
3.2 METODOLOGIA.....	80
3.2.1 Coleta de descrições de localização - <i>online</i>	82
3.2.2 Coleta de descrições de localização - presencial.....	86
3.2.3 Análise e desmembramento das descrições de localização e extração dos termos de relações espaciais	89
3.2.4 Proposição de Hipóteses para os testes <i>online</i> e em campo	90
3.2.5 Teste de associação livre de relações espaciais - similaridade	91
3.2.6 Análise, contagem e agrupamento inicial dos termos de relações espaciais ...	92
3.2.7 Elaboração de regras para aquisição de termos semanticamente similares....	92
3.2.8 Proposição de Hipóteses para o teste de associação livre	94
3.2.9 Agrupamento final dos termos semanticamente similares	95
4 RESULTADOS E ANÁLISES DOS RESULTADOS	96
4.1 COLETA DE DESCRIÇÕES DE LOCALIZAÇÃO - <i>ONLINE</i>	96

4.1.1 Coleta das descrições de localização e desmembramento em frases locativas.....	96
4.1.2 Extração e contagem dos termos de relações espaciais e seus respectivos <i>relata</i> e <i>locata</i>	98
4.1.3 Agrupamento e a classificação dos termos de relações espaciais.....	102
4.1.4 Novos termos utilizados como relações espaciais	107
4.1.5 Separação dos termos espaciais a partir dos mesmos <i>locata</i> e <i>relata</i>	111
4.2 COLETA DE DESCRIÇÕES DE LOCALIZAÇÃO - PRESENCIAL.....	115
4.2.1 Coleta das descrições de localização e desmembramento em expressões locativas	115
4.2.2 Extração e contagem dos termos de relações espaciais e seus respectivos <i>locata</i> e <i>relata</i>	116
4.2.3 Agrupamento e a classificação dos termos de relações espaciais.....	117
4.2.4 Novos termos utilizados como relações espaciais	120
4.2.5 Separação dos termos espaciais pelos mesmos <i>relata</i> e <i>locata</i>	123
4.3 COMPARAÇÃO E VERIFICAÇÃO DAS HIPÓTESES PROPOSTAS PARA OS TESTES <i>ONLINE</i> E PRESENCIAL	124
4.3.1 Comparação dos termos de relações espaciais nos dois testes	124
4.3.2 Comparação das classificações dos termos de relações espaciais encontrados com mesmos <i>relata</i> e <i>locata</i> dos dois testes.....	127
4.3.3 Verificação das hipóteses propostas para os dois testes	131
4.4 TESTE DE ASSOCIAÇÃO LIVRE DE RELAÇÕES ESPACIAIS – SIMILARIDADE.....	138
4.5 APLICAÇÃO DE REGRAS PARA CONTRIBUIR NA CRIAÇÃO DE AGRUPAMENTOS DE RELAÇÕES ESPACIAIS SEMANTICAMENTE SIMILARES.....	142
4.5.1 Primeira Regra	142
4.5.2 Segunda Regra	143
4.5.3 Terceira Regra	146
4.5.4 Quarta Regra.....	151
4.5.5 Quinta Regra	152
4.5.6 Agrupamentos finais.....	158
4.5.7 Distância entre os termos semanticamente similares	162

4.6 VERIFICAÇÃO DAS HIPÓTESES PROPOSTAS PARA O TESTE DE ASSOCIAÇÃO LIVRE	163
4.7 ESTRUTURAÇÃO DE EXPRESSÕES LOCATIVAS COM OS DADOS ADQUIRIDOS NOS TRÊS TESTES	170
5 CONCLUSÕES E RECOMENDAÇÕES	173
REFERÊNCIAS.....	176
APÊNDICE 1 – TERMOS DE CONSENTIMENTO DOS TESTES.....	191
APÊNDICE 2 – FOTOGRAFIAS DAS LOCALIDADES DO TESTE <i>OUTDOOR</i> ...	194

1 INTRODUÇÃO

O espaço geográfico pode ser descrito de diferentes maneiras, desde descrições por meio da linguagem falada, representações gráficas em forma de mapas e/ou representações computacionais. Na linguagem falada, assim como na cognição, a linguagem espacial torna-se uma fonte primária de informação. O fato de até as pessoas cegas, que não têm qualquer informação visual sobre o meio ambiente, serem capazes de navegar, indica que o espaço pode ser representado por diferentes tipos de informação e a informação linguística parece ser a mais útil neste caso (MAROTTA, 2010).

A linguagem se interpõe entre os humanos e o mundo, e também faz com que os humanos tenham acesso direto ao mundo. Desta maneira consegue-se pensar e visualizar como a linguagem natural¹ e o meio ambiente interagem. Sob esta perspectiva é que ocorrem as interações entre indivíduos pertencentes à mesma comunidade de fala, pois o simples fato de indivíduos se verem juntos em determinado espaço induz a uma interação: a linguagem (COUTO, 2007)

A interpretação da linguagem natural computacionalmente vem sendo estudada e tem se tornado essencial em áreas como a robótica (entendimento de comandos escritos ou de voz), inteligência artificial (processamento da linguagem natural - PLN), linguística computacional (extração de termos semânticos em textos), interação homem-computador (interfaces de programas de perguntas e respostas), *WEB* Semântica (ontologias), mineração de termos de relações espaciais e em Sistemas de Informações Geográficas (SIG). Em todos esses campos, foram desenvolvidos diferentes formalismos que abordam o problema de interpretação das expressões de linguagem natural em termos de situação ou significado contextualizado (KORDJAMSHIDI; OTTERLO; MOENS, 2010; MARCIŃCZUK; OLEKSY; WIECZOREK, 2016; REGIER, 1992). Principalmente na área da robótica e de SIG o avanço tecnológico fez crescer os estudos de representação das relações espaciais topológicas, como “dentro, fora, cruzar, começar e terminar”. Bruns e

¹ A expressão “linguagem natural” também pode ser conceituada como sinônimo de discurso comum, isto é, a linguagem usada habitualmente na fala e na escrita. A linguagem natural traz desafios para a sua representação computacional e estudos se iniciaram com processo de indexação, busca e recuperação de informações, como por exemplo, textos recuperados em bases de dados onde os termos do título e do resumo apresentavam sentenças em linguagem natural (LOPES, 2002).

Egenhofer (1996) apresentaram uma imposição de ordem para as relações serem representadas espacialmente, o que resultou numa aplicação formal de similaridades; Sharif, Egenhofer e Mark (1998) estudaram um modelo formal para a geometria de relações espaciais que distingue propriedades topológicas e métricas; Schwering e Raubal (2005) investigaram a integração de relações espaciais nas medidas de similaridade semântica entre conceitos geoespaciais e Cuayáhuitl et al. (2010) propuseram um sistema de diálogo na língua alemã, para gerar instruções de rota *indoor*.

Considerando as análises realizadas em SIG, pode-se entender a importância de realizar interpretações com relações espaciais vagas, como procurar informações relacionadas a locais imprecisamente definidos como “no sul” ou “no norte” de lugares. Além do mais seria interessante que os pedidos de informações em linguagem natural de texto fossem interpretados por estes programas (HALL; JONES, 2008; SHARIF; EGENHOFER; MARK, 1998). Portanto, programas de geoprocessamento e SIG podem ser implementados com estudos de outros tipos de relações espaciais, além das relações topológicas.

Porém, os estudos de relações espaciais se desenvolveram em simulações de locais restritos (ambientes fechados) onde eram dispostos alguns móveis ou objetos em diferentes posições, ou na tela de um computador, ou através de desenhos (CARLSON; LOGAN, 2001; FEIST, 2004; MUNNICH; LANDAU; DOSHER, 2001). Estas simulações continham limitações, pois foram testadas em computadores e pouco se enquadravam em situações reais vividas pelas pessoas em seu cotidiano, seja no ambiente *indoor* que se caracteriza por edificações que possuem pouco ou nenhum contato com o ambiente externo (*shopping centers*, hospitais, edifícios comerciais, industriais ou educacionais) (VIAENE et al., 2014); quanto em ambiente *outdoor* que são ambientes ao ar livre podendo ser urbano ou rural (ruas, praças, parques). Isto porque situações espaciais no mundo real, no espaço geográfico são mais complexas do que tarefas simuladas (HIMMELBERGER, 2015). Pessoas, no seu cotidiano, respondem naturalmente a perguntas do tipo “onde”, e este conhecimento deve ser obtido em testes com seres humanos e não baseados somente em dados geométricos ou topológicos (WINTER et al., 2018). Além disso, a diversidade de informações dispostas nos ambientes do mundo real (feições geográficas e pontos de referência) faz com que os usuários

apresentem, na sua linguagem natural falada/escrita, descrições diferentes de uma mesma região, utilizando mais de uma opção de relações espaciais.

Em um exemplo de estudo de aplicabilidade de relação espacial para a robótica, Kurata e Shi (2008) exploraram dois modelos formais (expressões relacionadas a direções e conjuntos específicos de expressões relacionadas à topologia) de relações espaciais para interpretar algumas expressões típicas que as pessoas usam para descrever uma rota. Segundo os autores, a utilização de uma ontologia espacial permitiria que robôs móveis navegassem em espaços humanos e interpretassem uma grande variedade de expressões de rota. Pois, uma ontologia de relações espaciais além de trazer conceitos pode trazer termos semanticamente similares para serem aplicados computacionalmente.

Blaylock, Swain e Allen (2009) mostraram que no contexto da geoinformação existem estudos que tentam utilizar a compreensão da linguagem natural para extrair referências de entidades geoespaciais (como empresas, ruas, interseções, parques, etc.) que são pontos de referência usados para formular consultas em BDG, mas não as relações espaciais envolvidas. Os autores afirmam que a busca e a extração de pontos de referência são processos que produzem resultados incertos. Além disso, explicaram que na *WEB* existe uma grande quantidade de informações geoespaciais em documentos, transmissões de áudio e *blogs* que poderiam ser codificadas em descrições qualitativas em linguagem natural (BLAYLOCK; SWAIN; ALLEN, 2009).

Considerando estas questões, o estudo das relações espaciais utilizadas em descrições de localização em linguagem natural torna-se um desafio porque as pessoas se expressam de maneiras diferentes. O ser humano utiliza a linguagem natural para se expressar dependendo de suas experiências espaciais, por exemplo, a percepção, navegação e representação do espaço, além do reconhecimento de objetos e de como pretende fazer a busca de objetos em um ambiente. Enfim, o que diferencia o ser humano das outras espécies é sua capacidade de utilizar uma linguagem para se expressar (LANDAU; JACKENDOFF, 1993; MACHADO, 2001). Em descrições em linguagem natural falada/escrita, as relações espaciais podem ser expressas através de uma variedade de termos e frases, o que dificulta a implementação destes termos em sistemas de localização espacial que tentam utilizar a linguagem natural. Estas relações espaciais não são independentes umas

das outras, e uma variedade de relações podem ocorrer e interagir numa mesma descrição. A maioria dos termos espaciais, incluindo aqueles que são utilizados no dia a dia das pessoas em diversas situações possui uma semântica muito variada, o que à primeira vista, não é perceptível (FERREIRA; DELAZARI, 2019b; WANG et al., 2008).

Percebendo a relevância das relações espaciais, as pesquisas sobre este tema visam entender como as pessoas os utilizam no seu cotidiano realizando descrições espaciais. E, dentre a diversidade de termos que podem ser utilizados como relações espaciais (em frente à/ao, no/a, perto da/do/de, próximo à/ao, ao lado da/do/de, entre, ao sul/norte/leste/oeste, à direita/esquerda, atrás, ao redor) muitos são utilizados como sinônimos/similares na linguagem natural falada, como no caso de “próximo” e “perto”. Uma razão para a ambiguidade das palavras espaciais é que muitos termos ficam claros em uma conversa presencial, em que o falante e o ouvinte estão no mesmo ambiente e têm consciência a qualquer momento daquilo que um ou outro quer expressar. A relação das palavras com os pensamentos e a relação das palavras com outras questões humanas, como por exemplo: a realidade, a comunidade, as emoções e as relações sociais podem ser tratadas pela semântica (PINKER; RAVAGNANI, 2008).

De acordo com o “Michaelis Moderno Dicionário Inglês & Português” (2021) a palavra semântica é um substantivo feminino, que para a filosofia: é a ciência que estuda a evolução do significado das palavras, signos e símbolos que estão a serviço da comunicação, semiologia; e para a linguística: 1 - é um ramo da linguística que estuda a significação das palavras e suas mudanças de sentido ao longo do tempo, bem como a representação do sentido dos enunciados; e 2 – é o significado dos vocábulos, por oposição à sua forma. E a palavra similaridade é um substantivo feminino que expressa a qualidade, caráter ou condições do que é similar. Por isto que os autores Moldovan et al. (2015) explicaram que os estudos de similaridade semântica devem ser realizados na língua em questão e não a partir de bases de dados em línguas diferentes, pois existem diferenças nas variáveis ortográficas e na semântica das línguas.

Rodriguez e Egenhofer (2003) explicam que a similaridade semântica é uma função da distância entre termos numa estrutura hierárquica de ontologias. Freitas (2007) explica que a similaridade semântica pode comparar elementos de duas ontologias através dos significados dos mesmos, buscando sinônimos ou outras

relações semânticas entre os elementos. Computacionalmente, as ontologias são essenciais para a descoberta semântica, pois a partir delas são feitas as análises entre os parâmetros existentes e as requisições de usuários (SENA et al., 2010). Algumas áreas que aplicam ontologias são: a biomedicina (PESQUITA et al., 2009), *WEB Semântica* (YUE et al., 2011) e as geociências. No contexto das geociências, entre as suas aplicações estão as Infraestruturas de Dados Espaciais (IDE) (STOCK et al., 2012), elementos relacionados a pontos de referência e feições geográficas (VARANKA, 2011), metalinguagem semântica (STOCK et al., 2013), ontologia de domínio das divisões e limites geográficos territoriais do Brasil (MARCH CASTAÑEDA, 2017), na criação de um catálogo de serviços geoespaciais baseado em semântica (YUE et al., 2011), em nomes de localidades chamados de textos geográficos colhidos em páginas da *WEB* e medidos através dos contextos encontrados (HAMZEI; WINTER; TOMKO, 2019; HU, 2018).

Dentro do contexto espacial, as pessoas utilizam a linguagem natural falada/escrita para descrever posições de lugares, objetos, outras pessoas, feições geográficas, pontos de referência, enfim fazem descrições de localização com muita frequência utilizando termos de relações espaciais diversos e não somente os termos topológicos. Além disto, a análise destas descrições faladas em linguagem natural pode fornecer o entendimento de como as pessoas utilizam as relações espaciais na determinação de posições/localizações no espaço em que se encontram inseridas, tanto em ambientes *indoor*, como em ambientes *outdoor*. Contudo, o (re)conhecimento dos ambientes do ponto de vista do cotidiano das pessoas é limitado, basta ver a grande quantidade de usuários que consultam mapas, aplicativos de rotas ou solicitam descrições verbais/textuais de localização; além do mais, as descrições podem conter implicações diferentes para ambientes *outdoor* e *indoor* (TENBRINK; BERGMANN; KONIECZNY, 2011). Podem-se considerar também situações nas quais as pessoas precisam informar sua localização para um possível resgate de emergência causado por um acidente ou por perda de localização em lugares desconhecidos.

Descrições em linguagem natural em diversas línguas que retratam o cotidiano das pessoas são de grande valia para examinar a linguagem espacial e seus termos, como por exemplo, *Cross the river* (Atravesse/cruze o rio) ou *The route is about 2 km* (O percurso tem aproximadamente 2km). Porém até agora foram pouco exploradas (EGOROVA, 2018). Estudos em outras línguas como é o caso do

alemão (BOSSE; PAPAFRAGOU, 2010), do italiano (URSINI, 2017), do inglês e do mandarim (LI; HUANG, 2020), analisaram algumas preposições de relações espaciais e não a sua totalidade.

Em descrições de localização em linguagem natural, tanto falada quanto escrita, partindo-se da pergunta “Onde você está localizado/a?” Podem surgir respostas diferentes, como por exemplo: Uma pessoa poderia responder: “Estou localizada em frente a um posto de gasolina.”, enquanto a outra pessoa poderia responder: “Estou localizada perto de um posto de gasolina.” Percebe-se a utilização de duas relações espaciais para a representação da mesma localização. Neste contexto, o problema da pesquisa se dá pela pergunta: A partir da obtenção dos termos de relações espaciais mais utilizados na linguagem natural do português brasileiro, como agrupá-los e classificá-los como semanticamente similares a ponto de serem entendidos e utilizados computacionalmente?

Para tanto, esta pesquisa parte da hipótese de que é possível agrupar os termos de relações espaciais descritas em linguagem natural por meio de um conjunto reduzido de termos semanticamente similares a partir de ambientes reais do cotidiano das pessoas. E, como consequência será possível estabelecer regras que contribuam e facilitem o processo de identificação e análise dos grupos de relações espaciais utilizados como semanticamente similares. Espera-se que tais regras possam ser aplicadas para outras línguas.

A pesquisa faz parte do projeto “*Where am I? Os Desafios da Interpretação de Descrições em Linguagem Natural de Localizações Geográficas*” em desenvolvimento através de uma parceria entre o Programa de Pós-Graduação em Ciências Geodésicas da UFPR e o *Institute of Natural and Mathematical Sciences* da *Massey University* de Auckland na Nova Zelândia. Este projeto trata da importância de como determinar a localização/posição do usuário com base em descrições em linguagem natural. Entende-se que existem situações nas quais o usuário precisa informar sua localização, porém não conhece o local onde se encontra, ou não dispõe de acesso ao posicionamento GNSS ou localização via celular. Dentro do contexto do projeto *Where am I?* onde esta tese terá a sua colaboração, está a tarefa de elaborar um conjunto de interfaces que possibilite a transformação de descrições de localização para um posicionamento em uma base cartográfica.

Além disso, esta pesquisa contribuirá com a descoberta dos termos semanticamente similares de relações espaciais para serem inseridos no conjunto de interfaces do projeto citado acima. Também contribuirá para aplicações computacionais que utilizem a linguagem natural no contexto espacial, bem como programas de PLN.

O primeiro trabalho nesta linha foi desenvolvido por Ferreira (2019) que investigou a influência do contexto na escolha e na forma de aplicação dos termos de relações espaciais.

1.1 JUSTIFICATIVA

De acordo com Hatzivassiloglou e Wiebe (2000) ao longo do tempo, com o crescimento, expansão, utilização local e remota da informação armazenada em grandes sistemas de recuperação da informação, muitos autores se dedicaram ao estudo das linguagens naturais e controladas aplicadas à indexação, extração e recuperação de informações. Lopes (2002) mostra em sua revisão de literatura, que a linguagem natural vem sendo estudada desde o início da década de 1970 para estratégia de busca em um ambiente de base de dados, bem como suas vantagens e desvantagens. Porém o estudo dos termos que compõem descrições em linguagem natural demanda de exploração e análises para que sejam aplicados computacionalmente. As ferramentas do PLN ou as aplicações em geoprocessamento tenderão a ser cada vez melhores a partir da aplicabilidade das descobertas realizadas a partir de descrições de localização, ou seja, o estudo da linguagem natural tem mais potencial do que somente indexação, recuperação ou mineração de textos.

Os tratamentos da linguagem natural devem considerar os componentes morfológicos, lexicais, sintáticos, semânticos e lógicos que intervêm na linguagem, porém, ainda não atingiram o estado de representação conveniente da linguagem natural (DE BRITO, 1992; KORDJAMSHIDI et al., 2013). Dentro do contexto do tratamento da linguagem natural falada, com relação ao componente semântico, deve-se levar em conta a cognição humana e a percepção do ambiente. Observadores com diferentes experiências realizam rotineiramente descrições de localização de feições geográficas que são desafiadoras de serem tratadas

computacionalmente para que sejam identificadas suas posições e suas orientações relativas. Estes fatores dificultam a aplicabilidade da linguagem natural para o ambiente computacional, onde também existe a dificuldade da conversão da linguagem natural falada para que resulte numa localização em uma base cartográfica. Uma solução para se buscar num Banco de Dados Geográfico (BDG) situações que tragam ou refinam uma localização solicitada, poderia ser realizada através de um pré-processamento das descrições envolvidas no processo. Uma possível conversão de entradas textuais em um conjunto de atributos, por exemplo: *locatum*², relações espaciais e *relatum*³, para que desta maneira sejam obtidos pares conectados conjuntivamente, sendo todos verdadeiros e que possam ser logicamente combinados (CHEN et al., 2018; STOCK; DELAZARI, 2011).

Ainda existem questionamentos com relação a metodologia adotada em estudos com linguagem natural. Wang et al. (2008) trabalharam no contexto de recuperação de imagens na língua inglesa e mostraram fraquezas em métodos onde indivíduos deveriam responder “sim” ou “não” para uma série de perguntas sobre uma determinada configuração e deveriam classificar uma lista de relações espaciais que melhor descrevessem uma determinada configuração ou métodos que tratassem somente das preposições espaciais deixando de lado a descrição de relações espaciais mais complexas ou até incomuns. Os autores afirmaram que as relações espaciais não são independentes umas das outras, porque uma variedade de relações pode ocorrer e interagir em uma descrição em linguagens naturais. O que realmente interessava eram quais relações os usuários se referiam ao descreverem uma imagem e como eles as utilizavam.

Na utilização de uma linguagem existem os conjuntos de conceitos das palavras, e a combinação destes conceitos em uma frase é chamada de estrutura sintática. Linguistas chamam o conjunto de conceitos e os esquemas que os combinam de “semântica conceitual” que é a linguagem do pensamento. Esta linguagem permite enquadrar uma situação de maneiras diferentes através da similaridade semântica, e muitas vezes até incompatíveis entre si. Dentro do contexto espacial, a linguagem é definida não apenas pela geometria dos objetos

² *Locatum*: pessoa, objeto ou evento, que pode ser estático ou dinâmico, cuja localização é descrita (KORDJAMSHIDI et al., 2012)..

³ *Relatum*: a entidade de referência (objeto, ponto de referência) em relação à qual a localização ou a trajetória é especificada (KORDJAMSHIDI et al., 2012).

e/ou feições geográficas, mas pelo modo como as pessoas a utilizam e invocam um sistema de eixos (PINKER; RAVAGNANI, 2008).

Stock (2008) explica que historicamente existem dois grupos de métodos para determinar a similaridade semântica entre conceitos ou comportamento: aqueles que usam alguma representação explícita de semântica (por exemplo, uma ontologia), e aqueles que usam a natureza implícita do objeto que está sendo representado (o esquema de um banco de dados com seus conjuntos de dados). O último grupo de métodos não mostrou ser muito eficiente e está sendo, em grande parte, abandonado. O primeiro grupo de métodos utiliza uma representação explícita, isto é, uma representação semântica. A autora em seu estudo de utilização da semântica em metalinguagem mostra que os métodos mais utilizados para representar a semântica de serviços da *WEB* e os objetivos dos usuários (isto é, informações comportamentais, funcionais ou processuais) envolvem ontologias de serviços *WEB*.

Stock (2014) explica que ontologias de relações espaciais têm sido definidas na literatura, mas a maioria está limitada a um pequeno grupo de relações, centradas em expressões de linguagem e não estão geometricamente especificadas. Isso mostra a necessidade de desenvolver um mecanismo para transladar relações espaciais não topológicas em perguntas que possam ser executadas para a representação num sistema métrico.

Dentre as classificações de relações espaciais mais comuns na literatura, as mais estudadas são as relações espaciais topológicas. Clementini e Di Felice (1997) explicam que a característica comum dos modelos de relações topológicas, é que eles fornecem uma base computacional para o raciocínio espacial. Portanto, estes modelos são uma compensação entre o formalismo necessário para um sistema de informação e a percepção humana do espaço geográfico. Apesar de ser extensamente estudado em todas as suas implicações teóricas e até mesmo implementados em programas de SIG, os modelos de relações topológicas sofrem a seguinte crítica: "Este é um modelo bom, mas é útil para uma verdadeira aplicação geográfica?" A crítica surge porque os atores desses modelos são entidades espaciais simples da geometria euclidiana, enquanto a realidade geográfica é permeada por entidades intrinsecamente complexas. Uma grande limitação dos atuais modelos de dados SIG é o descompasso entre realidade geográfica complexa (não estruturada, com limites incertos) e modelagem geométrica (objetos simples

com limites nítidos). A maioria dos modelos que representam relações espaciais assumem objetos com limites precisos (CLEMENTINI; DI FELICE, 1997), o que na maioria dos casos não acontece nos ambientes da vida real.

Compreender descrições espaciais e extrair automaticamente dados espaciais em linguagem natural é um problema desafiador em robótica, navegação, e na interação homem-máquina. A extração da semântica espacial da linguagem natural tem como desafios: 1) a identificação de palavras que desempenham os papéis espaciais e seus conceitos e 2) a classificação destas palavras com relação à: a) pessoa, objeto ou evento, que pode ser estático ou dinâmico, cuja localização é descrita (*locatum*); b) a entidade de referência (objeto) em relação à qual a localização ou a trajetória é especificada (*relatum*); e c) a relação espacial entre o *relatum* e o *locatum* (KORDJAMSHIDI et al., 2012).

A enorme quantidade de informações textuais, na maioria das vezes não estruturada, porque são produzidas pelos humanos no seu dia a dia, não é diretamente processável computacionalmente. E, a análise automática de texto envolve uma profunda compreensão de linguagem natural por máquinas, uma realidade da qual ainda se está muito distante (CAMBRIA; WHITE, 2014), que pode ser minimizada com estudos e análises de termos que compõem descrições em linguagem natural.

Diariamente nos sistemas de busca da *WEB* são realizadas perguntas para localização e instruções de navegação. Estas perguntas são respondidas em termos da nomenclatura dos locais que se querem localizar e é apresentada uma descrição de rota, quando solicitada. No entanto, estas respostas geradas nesta interação homem-computador são diferentes de um diálogo entre duas ou mais pessoas que trazem um contexto à informação. No futuro, são previstas ferramentas que forneçam respostas semelhantes às respostas realizadas em linguagem natural, não só a recuperação de documentos e trechos de textos, que na maioria das vezes trazem informações irrelevantes. Para que isto não ocorra, estudos em linguagem natural devem ser investigados como pré-requisitos importantes. Perguntas do tipo “onde” em contextos geográficos devem ser analisados através de descrições dos lugares. As descrições em linguagem natural são essenciais para a compreensão das respostas dadas pelos seres humanos quando questionados sobre lugares e suas localizações. Um exemplo de estudo de semântica para estas situações é apresentado através da captura de toponímia de estados, países, cidades, hotéis,

montanhas, entre outros, sem o estudo, trabalho e investigação das relações espaciais envolvidas (HAMZEI; WINTER; TOMKO, 2019).

A interação de dados espaciais requer sistemas que compreendam a linguagem natural espacial e não apenas processem as referências locais (DERUNGS; PURVES, 2016). Porém, qualquer sistema computacional que trabalhe com dados espaciais ou interaja com o ambiente deve ter a capacidade de se comunicar com os seres humanos sobre os ambientes. A comunicação espacial homem-máquina por meio da linguagem natural apresenta grandes problemas com as relações espaciais em linguagem natural porque elas são altamente ambíguas e com termos semânticos vagos. Isto dificulta a formalização destes termos para serem utilizados computacionalmente. Pois, se um computador fosse capaz de entender e expressar configurações espaciais em linguagem natural seria a chave para eliminar problemas de comunicação homem-máquina e aumentar a usabilidade dos sistemas (SCHWERING, 2007).

As abordagens no nível de palavras e conceitos em PLN são um primeiro passo para o entendimento da linguagem natural. O futuro do PNL reside em paradigmas computacionais que permitam a compreensão de uma narrativa, ou seja, a "criação do sentido". As técnicas da inteligência computacional devem ir além da representação sintática de documentos, devem imitar a maneira como o cérebro humano processa a linguagem natural, porém não será uma tarefa fácil por se tratar de cognição humana, onde cada indivíduo age e reage de acordo com sua percepção e conhecimento (CAMBRIA; WHITE, 2014). Não há regras nas descrições de localização em linguagem natural realizadas pelas pessoas em seu cotidiano. As descrições de localização em linguagem natural são realizadas com termos de relações espaciais e/ou com seus termos semanticamente similares, em sentenças muitas vezes mais complexas do que a estrutura das expressões locativas (*relatum*, relação espacial, *locatum*). Além disto, o grau de liberdade que a linguagem natural proporciona às pessoas traz desafios para a representação computacional dos termos de relações espaciais, principalmente porque se deve considerar que o uso indevido de um termo de relação espacial entre elementos pode transmitir informações enganosas de localização/posição (FERREIRA; DELAZARI, 2019b). Adicionalmente, cabe ressaltar que cada língua tem suas particularidades na utilização de termos de relações espaciais.

Dentro do contexto das relações espaciais, pode-se perceber que as relações topológicas tem recebido um maior tratamento nas duas últimas décadas na linguística e na matemática (CARLSON; LOGAN, 2001; HERNÁNDEZ-PASTOR; PERIÑÁN-PASCUAL; VALENCIA, 2016; KURATA; SHI, 2008; RICHTER; WINTER; RÜETSCHI, 2009) porque possibilitam representações espaciais formais dos elementos envolvidos nas relações (EGENHOFER; CLEMENTINI; DI FELICE, 1994; HALL; JONES, 2008; KURATA; SHI, 2008; RAGNI; TSEDEN; KNAUFF, 2007). De acordo com (VASARDANI et al., 2013), as pesquisas com termos de relações espaciais topológicas ainda estão sendo estudadas para o PLN. Porém, as relações espaciais também são classificadas em direcionais, de distância, de adjacência, projetivas, cardinais e são muito utilizadas na linguagem natural no cotidiano das pessoas, o que demanda mais pesquisas e estudos (CLEMENTINI, 2019; RODRIGUES; BENNETT, 2017; ZWARTS, 2005). No cotidiano as pessoas utilizam todos os tipos de relações espaciais. Descrições e sentenças de localização/posição de elementos, sejam eles quaisquer, utilizam termos que podem ser representados formalmente ou considerados semanticamente similares. Com os avanços tecnológicos as demandas por dados geográficos e cartográficos extrapolaram a necessidade de somente extrair ou recuperar informações. É necessário ir além à linguagem natural espacial principalmente no que diz respeito às relações espaciais e suas utilizações para que as aplicações computacionais tragam resultados o mais próximo possível da realidade.

Com relação aos termos semanticamente similares as pesquisas existentes são voltadas para os elementos, feições geográficas e localidades, sendo a grande maioria realizada na língua inglesa (BENNETT; AGARWAL, 2007; BLAYLOCK; SWAIN; ALLEN, 2009; CAMBRIA; WHITE, 2014; HAMZEI; WINTER; TOMKO, 2019; HAN et al., 2013; PINKER; RAVAGNANI, 2008). Portanto, analisar grupos de termos de relações espaciais encontrados em experimentos em linguagem natural torna-se importante, juntamente com a aplicação da similaridade destes termos, que é essencial para o PLN, auxiliar em sistemas de busca de informações na *WEB*, em aplicativos para localização e até em melhorias para os programas de geoprocessamento (HU, 2018). Pode-se justificar a importância deste estudo para a criação de uma ontologia de relações espaciais que possa ser utilizada para fazer a conexão de ontologias existentes e também em tarefas que requeiram a capacidade de inferência espacial (MARCH CASTAÑEDA, 2017).

Neste contexto, levanta-se a questão sobre a definição e estruturação de termos de relações espaciais semanticamente similares que possam ser utilizados em aplicações computacionais, tanto em programas de PLN como em programas de geoprocessamento e também, como contribuição para o projeto *Where am I?* - Projeto em desenvolvimento em conjunto com a *Massey University* da Nova Zelândia.

1.2 OBJETIVO GERAL

O objetivo geral desta tese é a proposição de uma metodologia para criação de grupos semanticamente similares de termos de relações espaciais utilizados em linguagem natural do português brasileiro que possam ser utilizados em aplicações computacionais.

1.3 OBJETIVOS ESPECÍFICOS

Tendo em vista o Objetivo Geral acima descrito, os objetivos específicos são:

- a) Planejamento de coleta de descrições de localização em linguagem natural em dois contextos diferentes e análise destas descrições para extração dos termos de relações espaciais mais expressivos, ou seja, mais citados pelos usuários;
- b) Realização de um teste de associação livre, através de sentenças de localização utilizadas no dia a dia das pessoas, a partir dos termos de relações espaciais mais citados nos dois contextos coletados e analisados no primeiro objetivo;
- c) Análise dos resultados do item “b” visando a formação de um agrupamento prévio dos termos de relações espaciais semanticamente similares;
- d) Elaboração e aplicação de regras que contribuam com a criação de um agrupamento de relações espaciais semanticamente similares que possam ser utilizadas computacionalmente.

1.4 ESTRUTURA E CONTEÚDO

A presente pesquisa está estruturada em cinco capítulos, sendo o primeiro relativo à introdução, a justificativa e os objetivos. O segundo capítulo traz a revisão de literatura dos cinco aspectos considerados básicos para a realização desta pesquisa: linguagem natural e espacial, relações espaciais, semântica, similaridade semântica e ontologia. O terceiro capítulo apresenta os materiais e a metodologia aplicada para o desenvolvimento da pesquisa, o quarto capítulo trata dos resultados e das análises. Primeiramente serão apresentados os resultados dos dois primeiros testes, suas análises, sua comparação e a verificação das hipóteses criadas para estes dois testes. Em seguida serão apresentados os resultados, a aplicação das regras criadas como consequência da criação do terceiro teste e a verificação das hipóteses para este teste. O quinto capítulo traz as conclusões e as recomendações futuras. Por fim, encontram-se as referências bibliográficas utilizadas no decorrer da pesquisa e os apêndices necessários para a realização da metodologia. As tabelas mais complexas foram adicionadas na base de dados científicos da UFPR e estão com seus respectivos *links* de acesso.

2 REVISÃO DE LITERATURA

2.1 LINGUAGENS NATURAL E ESPACIAL

Partindo-se do princípio que as crianças nascem em um mundo com gravidade, dotadas dos sentidos da visão e audição, e postura ereta, elas conseguem desenvolver uma percepção do espaço com propriedades específicas. Com isto, as crianças aprendem a aplicar termos espaciais utilizando sua linguagem natural para o espaço perceptivo onde se desenvolvem. As crianças sabem sobre o espaço e o tempo antes mesmo delas aprenderem em sua linguagem natural os termos relacionados a isto, portanto, a aquisição desses termos é construída em seu desenvolvimento cognitivo prévio (CLARK, 1973).

Uma linguagem se constrói através de um conjunto de símbolos cujos significados devem ser coletivamente conhecidos, identificáveis e traduzíveis, porque já foram, são e serão trabalhados em diversos contextos. Desta maneira, a linguagem tornou-se histórica e cultural para comunidades do mundo todo. A cultura das linguagens constitui-se através de processos pelos quais os indivíduos se identificaram como grupos, construindo comportamentos, valores, costumes, tradições que foram partilhadas, aceitas e assimiladas no decorrer dos tempos de geração em geração. A linguagem natural que possibilita esta troca de experiências entre as comunidades, de modo que estes processos sejam codificados e entendidos. Assim, o homem desenvolveu sua vida dentro de comunidades, que devem ser mantidas e preservadas para o seu próprio bem e para sua continuidade (JACKMAN, 1996; MACHADO, 2001). Ou seja, as línguas são essencialmente o resultado de uma interação cognitiva e vivencial entre o ser humano, o meio em que está inserido e a comunidade a que pertence. E também significam, acima de tudo, o que é experimentado no dia a dia. Não é só a mente que aprende a língua, o corpo também está presente em todo o processo. Por isso se pode dizer que, semanticamente, todo o sentido linguístico teve de ser sentido e cognitivamente experimentado pelo homem. Em outras palavras, o signo é apenas uma abstração teórica do sentido que as línguas comportam, portanto, nenhum falante tem acesso ao significado, mas todos usam e experimentam o sentido. Ao se encarar a língua como uma estrutura dependente da cognição humana, respectivas crenças e experiências físico-perceptivas, então os sentidos, mais do que os significados, não

são totalmente arbitrários, mas profundamente dependentes da relação cognitiva existente entre o falante e o mundo (físico ou mentalmente construído) ao qual a língua se refere (TEIXEIRA, 2006).

Clark (1973) explica que o homem é um habitante de um mundo que consiste em objetos, pessoas, espaço e tempo e por causa de sua composição biológica, ele percebe esses objetos, pessoas, espaço, tempo, e suas inter-relações de um modo particular. Estas são as condições sob as quais todos os falantes de qualquer língua nativa aprenderam a falar e descrever a localização dos objetos. O estado biológico e físico de cada indivíduo coloca um grande número de restrições prévias sobre como cada um descreve a localização dos objetos. Por exemplo, Físicos ou Matemáticos sentem-se inclinados em ver como a localização física pode ser especificada em geral, Geólogos ficam mais interessados em especificar a localização do homem e do ambiente com respeito a características terrestres, e Psicólogos se interessam em como o homem fala e interage com outras pessoas e que papel essas outras pessoas podem ter na especificação de seu espaço perceptivo.

O cérebro humano foi feito para nomear e rastrear coisas e objetos, isto pode ser percebido no vocabulário da maioria das línguas, onde se observa uma grande classe de substantivos e uma classe menor de palavras ou morfemas que especificam caminhos e lugares. Na língua inglesa a distinção destas classes é radical, sendo talvez até dez mil substantivos e aproximadamente oitenta preposições espaciais. Porém, o espaço também pode ser codificado por verbos, advérbios, adjetivos e sufixos. Consequentemente, o ser humano utiliza a linguagem natural para descrever o espaço de um modo que a geometria não consegue representar (PINKER; RAVAGNANI, 2008).

A expressão linguagem natural também pode ser conceituada como sinônimo de discurso comum, isto é, a linguagem usada habitualmente na fala e na escrita. No contexto computacional, existem desafios para o reconhecimento da linguagem natural, sendo que esta vem sendo estudada no processo de indexação, busca e recuperação de informações. E, no contexto da recuperação têm-se como exemplo, textos recuperados nas bases de dados onde os termos do título e do resumo apresentam apenas algumas sentenças em linguagem natural (LOPES, 2002).

A linguagem espacial faz parte da linguagem natural e (MAROTTA, 2010) explica que o espaço e sua codificação representam um tópico clássico no domínio da interação complexa entre linguagem e cognição. A autora explica que desde as obras de FILLMORE (1975), TALMY (1983; 1985), LAKOFF (1987) e LANGACKER (1987), as categorias linguísticas que expressam funções espaciais, estáticas e dinâmicas, foram investigadas em várias linguagens naturais. As espécies animais, incluindo os seres humanos, fazem algum tipo de representação mental do espaço, ou seja, da localização de objetos, lugares, caminhos e rotas. E, essas representações derivam de capacidades cognitivas especiais.

Gentner et al. (2013) explicam que aprender a representar e raciocinar sobre o espaço é crucial para o desenvolvimento cognitivo das crianças. Grande parte das evidências dos efeitos da linguagem espacial na cognição espacial vem de pesquisas que sugerem que a aprendizagem da linguagem espacial é que convida as crianças a formar representações conceituais correspondentes ao espaço. De fato, aprender a linguagem espacial pode influenciar a cognição espacial das crianças, pois a linguagem fornece sistemas simbólicos que potencializam maneiras de representar e raciocinar sobre o mundo. Portanto, a linguagem falada fornece ferramentas cognitivas, então deve ser possível observar essa influência quando as crianças adquirem sua língua.

Marotta (2010) explica que abordagens nativistas competem com as funcionalistas e relativistas no que diz respeito à cognição e a linguagem na expressão de categorias espaciais. E para tanto, existem duas hipóteses, sendo a primeira hipótese de que existe uma lista restrita de noções primitivas, universais e inatas topológicas, compartilhadas por todos os seres humanos, e codificado mais ou menos diretamente por adposições ou verbos. E a segunda hipótese, contrária a primeira, que as abordagens relativista e funcionalista assumem que a linguagem espacial é condicionada de várias maneiras e em vários graus por convenções culturais e refletem representações criadas pela exposição a palavras espaciais em relação a um idioma nativo. Ou seja, do ponto de vista linguístico, pode-se dizer que já se sabe quais são os elementos básicos que governam a estrutura do repertório léxico espacial encontrado através dos idiomas falados.

Pela linguagem espacial começa-se a examinar como os símbolos e as relações símbolo-visuais restringem o significado do mundo ao redor. Falar sobre o mundo espacialmente envolve falar sobre arranjos concretos de objetos, como

móveis em uma sala, objetos em uma imagem, cruzes e círculos em um diagrama, pontos de referência em determinados lugares e assim por diante. No contexto do aprendizado de uma segunda língua, a linguagem espacial também se apresenta como um desafio pelo significado e utilização dos termos espaciais e pelas suas particularidades que variam de língua para língua. Isto torna o aprendizado da linguagem espacial difícil na aquisição de uma segunda língua (COVENTRY, 2013).

As estruturas da linguagem espacial podem ser determinadas pela categorização pré-linguística do espaço, e pode existir a relação inversa, de tal forma que a estrutura e o léxico da linguagem espacial restrinjam a forma e as categorias do pensamento espacial. Em relação à experiência sensorial, existem duas hipóteses diferentes nas ciências cognitivas sobre a natureza das estruturas conceituais e representações semânticas: a Hipótese da Cognição Incorporada, segundo a qual o conteúdo é redutível à informação sensório-motora e a informação linguística desempenha um papel secundário; e a Hipótese do Conceito Abstrato, segundo a qual os conceitos são abstratos, e as entidades simbólicas não podem ser reduzidas a informações sensório-motoras, embora elas estejam relacionadas. A principal questão entre esses dois modelos é o tipo de representação semântica de termos lexicais (substantivos e verbos) referentes a entidades concretas e espacialmente determinadas. Na Hipótese da Cognição Incorporada, os conceitos são vistos como concretos e ancorados à experiência perceptiva do corpo humano; os conceitos abstratos são de alguma forma, secundários na representação cognitiva, e derivam de metáforas fundamentadas no sistema sensório-motor. Por outro lado, na Hipótese do Conceito Abstrato, as representações mentais podem ser de origem abstrata; e elas coletam e integram diferentes tipos de informação: linguística, perceptiva, sensório-motora e emocional (MAROTTA, 2010).

A linguagem espacial é estudada há décadas e é usada diariamente por indivíduos para descrever a localização de pessoas e objetos e para dar instruções de localização (HERSKOVITS, 1985; KORDJAMSHIDI et al., 2013; TENBRINK, 2017; WANG et al., 2008). As descrições de localização em linguagem natural podem ser simples ou complexas e podem incluir muitos termos e formas gramaticais diferentes (STOCK; YOUSAF, 2018). Para tanto, dentro da linguagem espacial podem ser consideradas duas classes: as preposições e os demonstrativos espaciais. Os parâmetros importantes para os demonstrativos espaciais também são importantes para as preposições espaciais e vice-versa; sustentar esta questão é

aceitar que um conjunto comum de componentes visuais e de ação possa sustentar uma linguagem espacial em geral. Estas classes envolvem relações subjacentes à sua utilização e relações sistemáticas com a forma de como o ser humano percebe e interage com o mundo, pois a linguagem espacial tem uma estreita correspondência com representações sensório-motoras dadas sua própria natureza, portanto o autor não assume que esses termos sejam abstratos desconectados da visão e da ação (COVENTRY, 2013).

Em muitas línguas a distinção espacial mais básica é entre algo que está próximo do falante e algo que está a certa distância, como os termos “aqui” e “lá”. Esta distinção é relativa e não absoluta, pois a expressão “coloque isto lá” tem sentidos diferentes para um operador de guindaste e para um neurocirurgião. Além disto, um quarto das línguas, incluindo o espanhol, faz uma distinção tripla, entre “perto de mim”, “longe de mim” e “no meio”; e bem poucas línguas, como a Tlingit (falada em Yukon) acrescentam mais uma distinção, “muito longe de mim”. Resumindo, nenhuma língua possui termos espaciais que meçam a distância em unidades. As distinções mais comuns nos idiomas do mundo são de sim ou não, ou se está dentro ou se está fora, literalmente. A questão não é somente que as línguas dividam o espaço em zonas com limites difusos, é que muitas das relações espaciais com as quais elas se preocupam são inerentemente qualitativas, que podem ser chamadas de distinções topológicas (PINKER; RAVAGNANI, 2008).

Uma das funções essenciais da linguagem é expressar as relações espaciais entre objetos e sua localização relativa no espaço (KORDJAMSHIDI et al., 2012). E Couto (2007) explica que as relações espaciais são um exemplo de que os conceitos mentais existem devido à existência de estímulos físicos do mundo externo. Dentro das relações espaciais podem-se citar as preposições quando foram proferidas por um indivíduo e dirigidas a outro numa comunidade; este processo teve início no cérebro do falante, onde foi associado um significado e foi transmitido pela fala aos ouvidos de outro falante. Por fim, esta interação foi possível porque os dois indivíduos estavam espacialmente próximos e no mesmo território. A interação através da linguagem começa como um processo psíquico e se torna social. O autor ainda cita que o estudo das relações espaciais de John Wilkins em 1668, quando representou relações em um modelo bidimensional com preposições de movimento, foi provavelmente a primeira tentativa na história dos estudos linguísticos de representação das relações espaciais.

As relações espaciais, entre objetos e pessoas e suas respectivas localizações relativas ou mudanças no espaço, são consideradas sob o ponto de vista da linguagem natural como a comunicação de informações espaciais. Diferenças gramaticais e características linguísticas semânticas desempenham papéis importantes na formação do significado das configurações espaciais. A principal tarefa na compreensão da linguagem espacial é identificar palavras ou frases que expressem a espacialização dentro de um determinado contexto (KORDJAMSHIDI et al., 2013; WANG et al., 2008).

Portanto, a percepção do espaço é uma das percepções mais básicas e estruturadoras de múltiplos modelos mentais que posteriormente o ser humano utiliza para dar forma aos modelos e processos com os quais conceitualiza e verbaliza as suas interações com o mundo (TEIXEIRA, 2013). Ou seja, as expressões linguísticas que representam o espaço dão pistas de como o mundo é percebido (KEWITZ, 2011).

Abordagens à linguagem espacial em linguística utilizam interpretações linguísticas e/ou imagens de relações espaciais (por esquemas), e estes são usados para representar palavras sem explicar processos da visão e de ação subjacentes a eles. Para a compreensão da linguagem espacial é necessário perceber como a linguagem e o mundo espacial são covariantes. A linguagem espacial pode ser representada por termos geométricos, termos que representem movimento e termos que representem o conhecimento da situação e/ou objetos envolvidos (COVENTRY, 2013; COVENTRY; GARROD, 2004).

2.1.1 Inglês, português e outras línguas

Existem estudos interlinguísticos em linguagem espacial principalmente comparando a língua inglesa a outras línguas. Estudos que abordam especificamente descrições de localização espacial entre o inglês e o português usado no Brasil são escassos. Tem-se conhecimento de um estudo que investiga o uso do termo de relação espacial *across*, que descobriu que em inglês esse termo é usado em objetos que são paralelos a uma superfície 2D e no português brasileiro é usado em uma superfície 3D correspondente em Inglês ao termo *through* (MCCLEARY; VIOTTI, 2004).

No entanto, alguns estudos interlinguísticos em português europeu e em inglês foram realizados. Batoréo (1998) enfoca o mapeamento do português europeu para a Metalinguagem da Semântica Natural de Wierzbicka (WIERZBICKA, 1972), destacando as diferenças entre as duas línguas. Também descreveu algumas diferenças no uso de verbos possessivos e existenciais; e nas conceitualizações que fundamentam algumas relações projetivas (especialmente na frente de). Batoréo e Duarte (2001) estudaram o uso dos verbos *ser*, *estar*, *haver* e *ter* no português brasileiro e os verbos *there is* e *have* em inglês.

Embora os estudos em português brasileiro sejam escassos, as preposições espaciais foram analisadas na língua Kriol (Guiné-Bissau) e comparadas ao inglês e ao português europeu, incluindo preposições simples [*a (to)*, *di (from)*, *na (in, at, on)*, *pa (para, por, em direção a)*, *sin (sem)*] e preposições complexas [*antis di (antes)*, *na banda di (perto de)*, *ba di (abaixo, abaixo)*, *dianti di (na frente de)*]. A preposição mais frequentemente usada neste estudo foi a preposição “na”. O objetivo do trabalho foi relacionar as preposições ao ambiente natural da linguagem na comunidade Kriol (COUTO, 2007).

Embora as comparações interlinguísticas do inglês e do português europeu sejam informativas, são insuficientes para fornecer uma compreensão completa das diferenças entre o inglês e o português brasileiro, já que as línguas evoluíram separadamente. No Brasil, o português tornou-se a língua oficial e foi influenciada por outras línguas ao longo dos séculos, tendo empréstimos de termos e vocábulos de uma ampla variedade de línguas que eram usadas no Brasil na época da colonização portuguesa (por exemplo, tupi Guarani), ou por imigrantes no Brasil incluindo escravos africanos e europeus (GUY, 1981; LIPSKI, 1975).

Estudos interlinguísticos de inglês e espanhol também foram realizados, incluindo (MARK; EGENHOFER, 1995), que descobriram que relações espaciais topológicas foram usadas de forma semelhante com estradas (linha) e parques (área) como *locatum* e *relatum* respectivamente, e mapeados para relações espaciais topológicas em proporções semelhantes nas duas línguas. Kemmerer (1999) analisou as diferenças na descrição de situações de localização “próximas” e “distantes” com demonstrativos pertencentes à classe dêitica das expressões, apontando que algumas línguas, como o inglês, faz uma distinção binária próximo/distante simples usando os termos: *this*, *here* e *that*, respectivamente; enquanto outros, como o espanhol, dividem o espaço em três ou mais setores, com

os termos: *este* (perto do locutor), *ese* (distância média do locutor) e *aquel* (longe do locutor). Além disso, um estudo de Bowerman (1996, p. 394) explicou que há diferenças no uso das preposições de contato, suporte e contenção em inglês, finlandês, holandês e espanhol. Por exemplo, o espanhol usa a preposição *en* para contato, suporte e contenção, mas também em situações mais especializadas usa *encima de* para suporte e *dentro de* para contenção. Os falantes de holandês usam três preposições diferentes: *op* para suporte, *in* para contenção e *aan* para contato; em finlandês, a palavra que termina com *-ssa* é usada para contenção e contato e com *-lla* para suporte.

Embora os estudos interlinguísticos de inglês e português brasileiro sejam escassos (RODRIGUES; SANTOS; LOPES, 2017), as duas línguas têm sido estudadas isoladamente. A linguagem espacial inglesa tem sido objeto de uma vasta gama de pesquisas, com particular atenção às preposições espaciais (COVENTRY; GARROD, 2004; FRANCO; ZAMPIERI; MENEGHELLO, 2013; HERSKOVITS, 1985; TYLER; EVANS, 2003), gramática e aos princípios que lhes são subjacentes, como *figure* e *ground*, quadros de referência e perspectivas (SHUSTERMAN; LI, 2016; TALMY, 1983, 2000; TENBRINK, 2017). Por outro lado, o português brasileiro é menos explorado, no entanto, o trabalho de Ferreira e Delazari (2019a) inclui uma análise de cinco termos de relações espaciais (*near*, *very near*, *next to*, *side by side* and *nearby*) e as distâncias que as pessoas aplicam a cada um destes termos; um estudo de descrições espaciais de movimento em alguns corpora dos séculos XVIII a XX em um dialeto específico do português brasileiro (português paulista) usando a classificação da linguagem de TALMY (1983) em línguas do tipo *verb-framed*⁴ e *satellite-framed*⁵ (KEWITZ, 2011); uma investigação das preposições “em” (*in*, *on*, *at*) e “para” (*to*) com os verbos “estar” (*to be*) e “ir” (*to go*) no sentido de movimento nas comunidades quilombolas de dois estados no Brasil (ARAÚJO, 2008); e uma análise semântica da preposição “em”, quando usada com artigos definidos em + a/o = na/no, e artigos indefinidos em + um/uma, demonstrativos em + este/esta ou

⁴ Línguas do tipo *Verb-framed* são as línguas que utilizam uma estrutura verbal para representar a relação espacial e o movimento, como o Português Europeu e o Francês (KEWITZ, 2011). São chamadas de Romance languages (TALMY, 2000).

⁵ Línguas do tipo *Satellite-framed* são as línguas que descrevem o movimento com uma preposição e a maneira é expressa através de verbos, por exemplo, a Língua Inglesa e a Alemã (KEWITZ, 2011). São chamadas de Germanic languages (TALMY, 2000).

aquele/aquela = neste/nesta, naquele/naquela, e todas estas combinações variando em número (plural ou singular) e gênero. A preposição “em” pode transmitir noções de contenção, suporte, contato e proximidade. Por exemplo, “O encontro será na escola, na Rua X, 12”. Sendo que a escola não se situa na superfície da rua, mas sim nas proximidades da rua (OLIVEIRA, 2012). Além disso, as preposições verticais “sobre” (*on*) e “sob” (*under*) podem depender de seis parâmetros diferentes: visibilidade, verticalidade, suporte, saliência, inclusão e proporcionalidade (RODRIGUES; GARCIA, 2015).

Em resumo, uma ampla gama de estudos interlinguísticos tem sido conduzida, porém focalizando um grupo limitado de termos de relação espacial, com atenção particular sendo dada às preposições espaciais e verbos de movimento. O inglês foi comparado a muitas outras línguas usando uma variedade de corpora e situações simuladas para estudar hipóteses específicas. No entanto, os estudos comparando Inglês e português brasileiro são limitados, incluindo o único estudo da preposição *across*. Embora haja mais trabalhos comparando o Português Europeu e outras línguas relacionadas, como o Espanhol com outras línguas (incluindo Inglês), o trabalho de Batoréo (1998) foi aplicado a alguns termos e suas diferenças. Além disso, embora o inglês tenha sido estudado com mais detalhes, existe um pequeno número de estudos de descrições de localização espacial em português brasileiro, e estes focam principalmente em um conjunto pequeno de termos, como no caso o trabalho de Ferreira e Delazari (2019a); de Rodrigues, Santos e Lopes (2017) na preposição “em” (*in, on, at*); e Rodrigues e Garcia (2015) nas preposições “sobre” (*on*) e “sob” (*under*).

Os estudos acima são de natureza teórica e exploram o uso de termos espaciais específicos em ambientes controlados com situações simuladas. Diferentemente do estudo de Fagundes et al. (no prelo) que aborda a comparação no uso de termos de relações espaciais do inglês (coletado por residentes da Nova Zelândia) e do português brasileiro (coletado por residentes no Brasil). O estudo analisou e comparou oito situações geográficas do mundo real (quatro no Brasil e quatro na Nova Zelândia), com a coleta de descrições de localização de ambos os grupos de entrevistados.

2.2 RELAÇÕES ESPACIAIS

Clark (1973) explica que os homens interagem socialmente uns com os outros e para que isto aconteça eles devem estar frente a frente e a uma curta distância. É nessa posição que a percepção se torna ótima para mensagens verbais e não verbais, pois se as pessoas estivessem lado a lado, ou de costas uma para a outra, ou uma de costas e outra de frente, ou em qualquer outra posição, não estariam em condições consideradas ótimas para interagir. Não é por acaso que conversas normais são realizadas cara a cara (frente a frente) e esta situação impõe particularidades nas expressões das relações espaciais.

As relações espaciais ultrapassam os aspectos puramente matemáticos quando são consideradas aplicáveis a problemas do mundo real, pois nesta situação devem ser considerados os aspectos linguísticos e cognitivos (EGENHOFER; HERRING, 1990). As pesquisas das áreas da cognição, psicologia e linguística estão interessadas em saber como os humanos percebem as relações espaciais entre feições geográficas, e seus estudos têm foco no uso das relações espaciais e suas conexões com a linguagem natural, enquanto pesquisas em cartografia e geografia estão preocupadas em apenas coletar os termos utilizados como relações espaciais (EGENHOFER; HERRING, 1990; MARK et al., 1995).

Teorias fundamentais de relações espaciais devem levar em conta a cognição espacial humana, pois pesquisas da área de cognição mostram teorias já publicadas ou mostram as intuições dos pesquisadores como base para formalização, e raramente desenvolvem conceitos utilizando experimentos com seres humanos (MARK et al., 1995). Além do mais, mostrar as distinções que os humanos fazem quando raciocinam e descrevem localizações utilizando relações espaciais em linguagem natural, mesmo adotando diferentes conceitos, seja pela cultura ou pelas características individuais (gênero, idade e grau de instrução) devem ser levadas em consideração no processo da análise das descrições em linguagem natural, para extração dos pontos de referência e das relações espaciais (MARK et al., 1995; PAPAFRAGOU; MASSEY; GLEITMAN, 2002). A cultura e o ambiente que levam a diferenças sobre quão prontamente uma ou outra capacidade mental é posta em uso, não a linguagem (PINKER; RAVAGNANI, 2008).

Tarefas do cotidiano das pessoas frequentemente requerem a utilização de relações espaciais, sejam descrições simples, como “O controle remoto está na

frente da televisão.”, ou em descrições mais complexas, como fazer a descrição da localização de um cinema em uma cidade quando alguém é questionado, ou explicar a localização de algum fato ocorrido, como um assalto ou um acidente. As relações espaciais vêm sendo estudadas em descrições e também para aplicações computacionais, desde a década de 1980 por psicólogos e linguistas (CARLSON-RADVANSKY; LOGAN, 1997).

As informações espaciais são entendidas e transmitidas através do uso de relações espaciais que descrevem como um objeto em uma cena ou uma imagem está localizado (*locatum*) em relação a algum outro objeto (*relatum*). As relações espaciais foram estudadas em disciplinas diferentes, incluindo ciência da computação, ciência da informação geográfica, ciência cognitiva e linguística. Dentro destas disciplinas, as relações espaciais são geralmente consideradas adequadas a uma das três categorias: topológica (sobreposições e separações), direcional (acima, abaixo, direita e esquerda) e de distância (perto e longe) (WANG et al., 2008). Ou seja, os termos de relações espaciais formais destas três categorias têm uma semântica definida em programas SIG, mas as relações espaciais em linguagem natural têm uma semântica complexa, muitas vezes implicando em mais de um tipo de relação espacial formal e/ou a utilização de mais tipos de relações espaciais (SCHWERING, 2007).

Egenhofer, Clementini e Di Felice (1994) definiram as relações espaciais topológicas com o desenvolvimento de definições formais, matemáticas e precisas. Uma ampla gama de pesquisas nas últimas duas décadas, trabalhou na ciência da informação geográfica com modelos matemáticos para relações espaciais topológicas. Além disto, ferramentas de programas comerciais, tanto de uso de especialistas como não especialistas, utilizam informações geográficas usando abstrações limitadas a pontos, linhas e áreas com bordas nítidas e bem definidas reduzindo as relações espaciais a funções de buffer (DERUNGS; PURVES, 2016). No entanto, as pessoas tendem a pensar em termos de semântica informal, através de modelos cognitivos de suas experiências. Esses modelos cognitivos podem diferir de pessoa para pessoa, pela sua cultura, pelo contexto, além de serem vagos. As abordagens formais de relações espaciais não conseguem representar esses tipos de variação. Essas abordagens geralmente requerem conhecimento especializado e compreensão de uma linguagem formal de definição e não coincidem bem com a semântica informal usada em contextos do cotidiano das pessoas (STOCK et al.,

2010). Como por exemplo, qual a relação espacial que melhor representa a FIGURA 1? A árvore está à esquerda, em frente, próxima, ou ao lado da casa? Todas estas relações podem ser utilizadas no cotidiano das pessoas e expressam a mesma situação geográfica.

FIGURA 1 - RELAÇÕES ESPACIAIS PARA UMA SITUAÇÃO GEOGRÁFICA



FONTE: A autora (2021).

As relações espaciais podem ser classificadas como topológicas, direcionais, projetivas ou cardinais, ou descrevem proximidade, adjacência ou orientação. As preposições topológicas são aquelas que descrevem aspectos como conectividade, inserção e contenção (por exemplo, dentro e fora). As preposições projetivas envolvem a descrição de uma relação relativa a algum eixo projetivo, que pode ser um eixo relativo ao observador, ao próprio objeto ou um quadro de referência externo (por exemplo, acima, atrás, além, na frente, à esquerda/direita de, abaixo, sobre). As preposições de proximidade descrevem a localização usando a distância qualitativa entre objetos (por exemplo, perto e perto) (COVENTRY; GARROD, 2004; KEMMERER, 2006; TENBRINK, 2017). Os termos de relação espacial direcional denotam movimento ou direção (por exemplo, através, ao longo, ao redor, sobre, através, em direção) (ZWARTS, 2017), enquanto aqueles que descrevem adjacência podem ser considerados como incorporando proximidade e projeção, embora nesta classe de relação espacial os termos não tenham sido bem explorados até agora (por exemplo, adjacente, junto, anexo). Os termos de relação espacial que descrevem a orientação descrevem os ângulos de objetos ou partes dos mesmos, em relação uns aos outros (por exemplo, de frente, paralelo, perpendicular, orientação) (BITTERS, 2009).

Embora tenham sido descritas as classificações acima, existem discordâncias entre os estudos existentes ou existem variações na composição das

classes. Pesquisadores propuseram diferentes esquemas de classificação para termos de relação espacial, incluindo Coventry e Garrod (2004), que definem relações direcionais, topológicas (com três subclasses diferentes: simples, de proximidade e pelo quadro de referência absoluto) e projetivas; Bitters (2009), que propõe uma série de classes, incluindo adjacência, proximidade, acima/abaixo, contenção, orientação e intersecção; Kemmerer (2006), que define classes dêiticas, topológicas, projetivas horizontais e projetivas verticais; Levinson, Meira e Group (2003), que propõem topologia, fixação, superposição e várias outras classes; Retz-Schmidt (1988), que distingue entre topológico e direcional/projetivo; Zwarts (2005), que definem as classes locativa e direcional; Zwarts e Winter (2000), que identificam relações direcionais e locativas, e então dividem as relações locativas em projetivas e não projetivas; e Tenbrink (2007), que distingue relações dimensionais/projetivas, topológicas, relacionadas ao percurso e à distância. Finalmente, Hois et al. (2009) e Bateman et al. (2010) identificam uma série de tipos de relações em sua ontologia GUM-Space, incluindo quantitativa e qualitativa, projetiva, topológica, distribuição, parcialidade. Esta tese tem como foco os termos de relação espacial da linguagem natural, ao invés dos esquemas que foram desenvolvidos para classes específicas de relações espaciais na literatura do Raciocínio Espacial Qualitativo baseado em conceitos teóricos, em vez da linguagem de uso cotidiano (COHN et al., 1997).

Couto (2007) explica que a fim de evitar mal-entendidos sobre os diversos usos de um lexema de uma língua dada, é importante considerar o contexto da situação, levando em consideração os lexemas que designam “elementos” e àqueles que se referem às relações entre esses elementos. A lexicalização de expressões pode ter a ver com a sobrevivência de nossos ancestrais. Por exemplo, na antiguidade era importante para os seres humanos saberem se o leão estava “atrás” ou “na frente” de alguma coisa. Do mesmo modo, era de vital importância saber se o leão estava “antes” ou “depois” de uma árvore ou colina, porque se estava “antes” estava próximo do observador sendo uma situação perigosa. Mesmo que o leão não estivesse por perto, era importante estar “dentro” de uma caverna se estivesse chovendo, trazendo a importância do conceito de “dentro” e “fora”, por exemplo. A preposição “no” é um bom exemplo, o observador está “no” quarto. Nesse caso, “no” pode indicar todas as posições possíveis: no teto, na parede e no chão; e no caso, a parede pode ser: a da frente, a de trás, a da direita ou a da esquerda do observador. O observador poderia usar também todas as preposições que codificam as relações

especiais específicas, ou seja, “sobre” para superioridade em vez de “no teto” e “embaixo” para inferioridade em vez de “no chão” e assim por diante. Desta maneira se confirma a relatividade das relações espaciais.

O entendimento dos termos de relações espaciais (em, em frente de, ao lado do, do outro lado, entre, à direita/esquerda, perto de) é importante por refletir a maneira como o ambiente observado é entendido e descrito pelos usuários por meio da linguagem natural. Diariamente as pessoas descrevem cenários e/ou localizações, ou seja, as relações espaciais apresentam como os elementos dos ambientes estão conectados entre si (FERREIRA; DELAZARI, 2019a), contribuindo para a melhoria dos ambientes.

Hall e Jones (2008) coletaram *relata*, *locata* e preposições espaciais de legendas de fotografias de locais isolados, na língua inglesa, em redes sociais e perceberam que o número de preposições utilizadas frequentemente como relações espaciais era pequeno. Na coleta dos dados, os autores identificaram as preposições espaciais, depois fizeram a procura por palavras que começavam com letras maiúsculas porque seriam os possíveis candidatos de *locata* e *relata*. Com este trio de informações conseguiram analisar quais ângulos e distâncias estavam envolvidos no uso de direções cardinais. Os autores continuam os trabalhos com fotografias nas escalas de ruas e pontos de interesse e com outras relações espaciais, sempre focando em relações espaciais vagas.

A mineração de palavras em *corpus* também tem interesse no estudo das relações espaciais, como foi o caso do estudo de 200.000 expressões que consideraram a busca da relação “próximo” em diferentes situações geográficas (escalas) entre dois lugares, ou seja, A próximo de B, onde A e B são topônimos. Desta maneira foram analisadas várias relações de proximidade em linguagem natural simples para geração de modelos para uso em sistemas que levariam em conta a imprecisão na linguagem. Na apresentação dos trabalhos relacionados com relações espaciais vagas, os autores trouxeram uma tabela com dezesseis estudos abrangendo os domínios da psicologia, linguística cognitiva, ciência da computação e SIG, entre os anos de 1978 a 2014 (DERUNGS; PURVES, 2016). Este resumo mostra a importância do estudo das relações espaciais para estas áreas, principalmente no caso das relações imprecisas e vagas.

2.2.1 Eixos de referência espacial

Clark (1973) explica que físicos e matemáticos provam que para se localizar um objeto no espaço, este objeto (*locatum*) deve sempre estar localizado em relação a outros objetos (*relatum*) no espaço. Por exemplo, a localização de um objeto pode ser exclusivamente definida no espaço tridimensional pela distância entre os objetos. Porém esta é uma maneira abstrata de especificar a localização e isso levou ao desenvolvimento de sistemas de referência, que tem como vantagem um sistema de coordenadas cartesiano familiar com seus eixos x, y e z que tornam as distâncias mais fáceis de definir e, uma vez definidas, mais fáceis de usar. As noções mais básicas de localização são: pontos de referência, linhas de referência, planos de referência e direções. Objetos estão localizados com relação a outros objetos no espaço tomados como: pontos de referência no contexto unidimensional, linhas de referência no contexto bidimensional e planos de referência no contexto tridimensional. Esses pontos, linhas e planos de referência, por geometria simples, definem as direções, tudo isto é fundamental para a concepção do espaço perceptivo do homem. Porém não são suficientes para a representação de situações vividas no dia a dia das pessoas.

Para o uso de relações espaciais o ideal é a seleção de um referencial, um conjunto de três eixos coordenados ortogonais cujo ponto de intersecção é a origem, desta maneira pode-se especificar se um objeto está “acima” ou “abaixo”, “na frente” ou “atrás”, “à esquerda” ou “à direita”. Uma direção também pode ser arbitrada. A formação do referencial ou o quadro de referência resulta em três situações:

- 1) Centrado no indivíduo, no observador ou egocêntrico: tem um eixo alinhado da cabeça aos pés, frente e costas, direita e esquerda e codifica a localização espacial dos objetos naquele momento e situação;
- 2) Centrado no objeto: é baseado nas propriedades físicas do objeto e usa um dos lados do objeto (superior, inferior, frontal, posterior, esquerda ou direita) para orientar os eixos, ou seja, é um referencial alocêntrico;
- 3) Centrado no ambiente propriamente dito, ou seja, referência geocêntrica: definem as relações espaciais em relação às

características do ambiente, como a direção da gravidade, a luz do sol, azimute, o campo magnético da terra. Ou seja, tem as propriedades do ambiente em questão para orientar seus eixos.

Estas três formações variam de observador para observador, portanto estes parâmetros não são fixos e a proximidade e a distância também afetam na escolha do referencial e conseqüentemente na seleção da preposição espacial a ser utilizada em uma descrição. Portanto, as disposições espaciais para os ambientes, sejam eles: internos, externos, do mundo real ou do ambiente computacional, são infinitas, mas as relações espaciais em qualquer linguagem natural têm um número limitado de palavras. Interessante que pela definição do sistema de referência aloccêntrico, um sistema de referência egocêntrico poderia ser considerado um caso especial do aloccêntrico, onde o objeto seria o próprio observador. Entretanto, a imagem corporal os diferencia, pois, o observador é dotado de movimentos que configuram a situação do corpo no momento da observação espacial que será realizada na maioria das vezes voltada para frente. Em contrapartida, os objetos e os pontos de referência são escolhidos pela confiabilidade enquanto referenciais espaciais e servem de base para outras localizações espaciais no sistema aloccêntrico (CARLSON-RADVANSKY; LOGAN, 1997; COVENTRY, 2013; MATSUSHIMA; RIBEIRO FILHO, 2003; MCNAMARA; RUMP; WERNER, 2003; PINKER; RAVAGNANI, 2008; TALMY, 1983).

Quando pessoas fazem descrições de formato ou de partes de um objeto móvel estão utilizando um referencial centrado no objeto, ou seja, um sistema de coordenadas que atravesse um objeto saliente, permitindo que suas partes ou que outros objetos sejam localizados em relação à sua frente, ou aos seus lados, ou em cima/embaixo. Porém, existem vantagens e desvantagens, embora seja útil para reconhecer formas e perceber organizações de objetos, um referencial centrado em objetos tem dificuldade para distinguir as duas dimensões horizontais de forma coerente. Alguns objetos possuem frente e trás naturalmente (bicicleta, televisão, geladeira), mas outros não, como por exemplo, árvores ou lustres. E para complicar, existem os lados direito e esquerdo que podem ou não serem facilmente distinguíveis, surgindo assim outro tipo de sistema de coordenadas: o referencial egocêntrico. Um problema deste referencial é que as pessoas saem do lugar, portanto o referencial é inútil para localizar coisas que não estejam ligadas às

peessoas, a menos que estas pessoas concordem em ficar em determinado local de frente para determinada direção. Outro problema é que o corpo e a mente humana são aparentemente simétricos, o que muitas vezes dificulta a distinção de direita e esquerda (PINKER; RAVAGNANI, 2008).

Além do ponto de observação e da posição dos objetos dentro de um determinado local em um determinado contexto, deve-se levar em conta a percepção do espaço e a dimensionalidade, por exemplo, para um espaço 3D as relações espaciais “atrás” e “em frente de” são apropriadas. Além desses fatores, diferenças individuais como gênero e lateralidade podem afetar a forma como se formam modelos mentais de fenômenos espaciais e os significados dos conceitos espaciais (WANG et al., 2008). Quando são utilizados lados dos elementos que podem ser definidos como “frente” ou “fundo”, como por exemplo, casas e carros a interpretação da relação espacial “atrás” é chamada “intrínseca”. Porém se a entidade de referência é uma árvore (pressupõe-se que não tem “frente” e “fundo”) e fala-se: “o carro está atrás da árvore”, entende-se que a árvore está localizada entre o carro e o observador, sendo denominada de interpretação “extrínseca” (BLÜHDORN, 2001).

Pesquisadores da cognição espacial acreditam que as pessoas (e até alguns animais) nascem dotadas da capacidade de usar os três referenciais (geocêntrico, egocêntrico e centrado no objeto) dependendo das tarefas e das circunstâncias. As pessoas interpretam o mundo nos referenciais disponibilizados pela linguagem, por exemplo, os falantes da língua inglesa são capazes de utilizar coordenadas geocêntricas muitas vezes utilizando os pontos cardeais; e muitos o fazem sem grande esforço, apontando para o norte dentro de uma sala por dedução, ou referenciando os locais dentro dos prédios por sua posição cardinal (salão nordeste, salão de conferência sul) (PINKER; RAVAGNANI, 2008). Porém no geral, as pessoas fazem um deslocamento de um ponto A para um ponto B baseadas no sistema de referência egocêntrico (MCNAMARA; RUMP; WERNER, 2003).

Para o espaço perceptivo do homem (eixo de referência egocêntrico), que tem postura bípede e é nesta posição canônica que realiza as atividades cotidianas, definem-se três planos de referência: um plano de simetria (o plano vertical que é simétrico biologicamente e separa a esquerda e a direita) e dois planos perpendiculares entre si sobre os quais a percepção é assimétrica (o plano vertical que separa a frente da parte de trás e o plano horizontal no nível do solo). A noção

de posição canônica tem várias consequências importantes para a linguagem, pois o homem move-se para frente, não para o lado ou para trás. Este fato se encaixa com a percepção do ser humano, que se move na direção sobre a qual ele percebe o ambiente. Define-se uma dimensão frente-trás na composição biológica do homem a partir da direção em que ele se movimenta. A vertical biológica e o nível do solo são ainda mais convenientes como planos de referência no espaço perceptível desde que o homem esteja em posição canônica, pois coincidem exatamente com a vertical do corpo humano e com o nível do solo geológico (CLARK, 1973).

As pessoas não possuem localizadores automáticos implantados em seus cérebros (como quando utilizam receptores GNSS para receber sinais de satélites) que forneçam suas localizações geográficas, por isto tem que eleger uma referência que possa ser identificada com segurança por outras pessoas (ou pela mesma pessoa em momentos diferentes) para então especificar a direção e a distância de um objeto em relação a ela. Para a dimensão para cima e para baixo, a gravidade é um referencial onipresente e constante. Mas as outras duas direções, norte/sul e leste/oeste, são problemáticas, porque não há bússolas ou quadros com a informação “você está aqui” espalhados pelo mundo para orientação. Para resolver este problema, uma opção é procurar um referencial geocêntrico: um eixo norte/sul ou leste/oeste alinhado a marcos territoriais, geográficos ou outra característica do local. O referencial geocêntrico tem a vantagem de ser fixo independentemente de onde se esteja, mas tem a desvantagem de ficar indisponível em ambientes fechados ou em ambientes desconhecidos e muitas vezes ser aplicado de maneira errada. Qualquer objeto ou parte que possa se mover manterá uma localização constante em relação a algum outro objeto a que esteja ligado, não em relação ao mundo (PINKER; RAVAGNANI, 2008).

A maneira como as relações espaciais são expressas e os referenciais aplicados variam entre as línguas. Por exemplo, em um dos idiomas australianos e no idioma maia são utilizados referenciais absolutos para se referir às posições relativas dos objetos, como por exemplo, só é possível localizar uma cadeira em relação a uma mesa em termos de pontos cardeais “A cadeira está ao norte da mesa”. E nestas línguas não se tem traduções como “A cadeira está na frente da mesa”. Em inglês, os quadros de referência absolutos são utilizados em escalas geográficas, por exemplo, uma cidade fica ao sul de outra (COYNE; SPROAT; HIRSCHBERG, 2010).

Em se tratando da classificação cognitiva do espaço, o mesmo pode ser classificado como: figura, vista, ambiental e geográfico. Esta classificação foi atribuída com base no tamanho projetivo do espaço em relação ao corpo humano, não em relação ao tamanho absoluto real ou aparente. Desta maneira o espaço pode ser percebido e ser cognitivamente tratado pela mente do ser humano a partir de sua visão. O espaço figural é projetivamente menor que o corpo, como é o caso de objetos e figuras. O espaço de vista é projetivamente maior que o corpo, como ambientes internos de casas, praças de cidades ou pequenos vales em ambientes rurais. Estes dois espaços são percebidos visualmente a partir de um único local, sem a necessidade de locomoção. O espaço ambiental é projetivamente muito maior que o corpo e o envolve, como bairros e cidades. Estes espaços necessitam de locomoção para serem percebidos e aprendidos cognitivamente, demandando períodos para serem captados. O espaço geográfico é projetivamente muito maior que o corpo e não pode ser percebido diretamente pela locomoção, ele deve ser aprendido por meio de representações simbólicas, como mapas ou modelos, que essencialmente reduzem o espaço geográfico ao espaço figurativo, onde acaba entrando o conceito das escalas de representação (MONTELLO, 1993).

A classificação do espaço tendo como base a cognição pode contribuir não somente nas experiências ambientais e suas representações em mapas, como também entender o porquê das escolhas individuais de diferentes quadros de referência usados para organizar o conhecimento espacial, e da capacidade espacial em diferentes escalas (MONTELLO, 1993).

2.2.2 Extração de relações espaciais

Pessoas no seu cotidiano são capazes de entender e reagir facilmente a descrições de lugares, mas computacionalmente, a extração de instruções adequadas das descrições exige o aprendizado de um conjunto de frases linguísticas que especifiquem automaticamente uma semântica espacial. O procedimento de aprendizagem pode explorar qualquer conjunto de características linguísticas que venha a ser extraído das frases e generalizado para frases não vistas. Esta técnica visa nomear informações espaciais disponíveis a partir de representações espaciais qualitativas, bem como informações em *corpus*

disponíveis a partir de dados linguísticos (KORDJAMSHIDI; OTTERLO; MOENS, 2010).

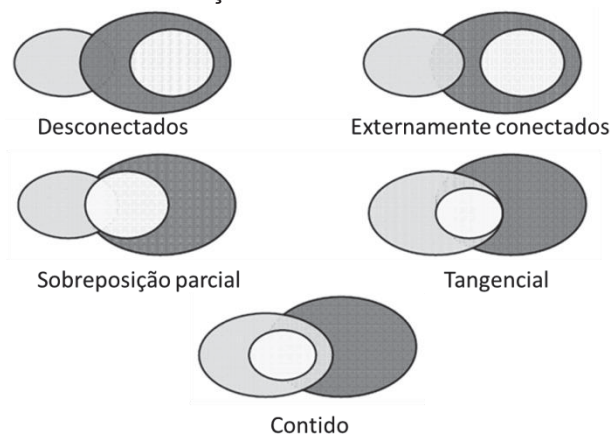
Os termos de relações espaciais desempenham um papel importante na compreensão da linguagem espacial, pois são cruciais nas descrições de localização no mundo todo. Um exemplo disto é o sistema “*WordsEye*” que converte automaticamente texto em linguagem natural em cenas 3D renderizadas. Como as relações espaciais são expressas explícita ou implicitamente, através de muitos termos em linguagem natural, para representar essas relações em cenas 3D torna-se necessário um conhecimento linguístico do idioma em questão e das propriedades espaciais e funcionais dos objetos. O modelo de expressão utilizado para a transformação das cenas 3D é *X-preposition-Y* onde o X é denominado *figure* e o Y é o *ground*, ou seja, para que o programa funcione os modelos devem estar em expressões locativas (COYNE; SPROAT; HIRSCHBERG, 2010).

Termos de relações espaciais são frequentemente preposições, particularmente na língua inglesa, embora também possam ser verbos, substantivos, adjetivos, advérbios, e demonstrativos (COVENTRY, 2013; KEMMERER, 1999; OLIVEIRA, 2013; PINKER; RAVAGNANI, 2008; STOCK; HALL, 2018). E a extração destes termos de relações espaciais da linguagem natural depende do contexto. Por exemplo, na frase *Give me the book on Artificial Intelligence on the big table behind the wall.* (Dê-me o livro sobre Inteligência Artificial que está na mesa grande atrás da parede.) a palavra “sobre” não é uma relação espacial porque especifica o título do livro, já as palavras “na” e “atrás”, são termos espaciais. Para este exemplo seriam extraídos dois trios de *locatum*, relação espacial e *relatum*: livro, na, mesa e mesa, atrás, parede. No caso de extrações automáticas de relações espaciais, mapear os termos que estão em contextos espaciais torna-se um desafio (KORDJAMSHIDI et al., 2012). A expressão locativa será mapeada e extraída, mas a verificação da espacialidade deve ser realizada por algum nativo da língua em questão.

Um experimento de raciocínio topológico realizado com nativos da língua Alemã e nativos da língua na Mongólia analisou a utilização dos termos “sobreposição” e “estar dentro” em expressões locativas, com o seguinte contexto: “Imagine que dois amigos estão falando sobre regiões coloridas em um mapa temático e um deles fala: A região branca está dentro da região mais escura. E, o outro responde: A região cinza se sobrepõe à região mais escura.” Estas duas expressões locativas trazem cinco possibilidades de como as três áreas podem ser

localizadas no mapa: desconectados, conectados externamente, sobreposição parcial, a primeira região pode fazer parte da segunda com as bordas se tocando (tangencial), ou a primeira região pode conter a segunda região, conforme ilustra a FIGURA 2. As pessoas utilizam expressões espaciais que descrevem as relações topológicas entre elementos, mas não comunicam as informações métricas entre os elementos. Estas são situações típicas no cotidiano das pessoas. Outra situação é quando as pessoas utilizam o termo “ao lado” sem especificar se o elemento citado está à direita ou à esquerda. Os resultados encontrados no experimento mostraram que as relações espaciais utilizadas pelos voluntários estavam relacionadas com os aspectos culturais e que pessoas com diferentes idiomas podem pensar de formas diferentes (RAGNI; TSEDEN; KNAUFF, 2007).

FIGURA 2 - MODELOS PARA AS RELAÇÕES ESPACIAIS “SOBREPOSIÇÃO” E “ESTAR DENTRO”



FONTE: Adaptado de Ragni, Tseden, Knauff (2007).

Portanto, para no tratamento de descrições de localização para extração de relações espaciais, faz-se necessário o desmembramento das descrições em expressões locativas, que contém pelo menos uma relação espacial e os objetos que permitam que esta relação exista (HERSKOVITS, 1985; LIU; VASARDANI; BALDWIN, 2014). Ou seja, tais expressões consistem no objeto cuja localização está sendo descrita (conhecido como *locatum*, figura, trajetória ou objeto localizado), outro objeto de referência (conhecido como *relatum*, solo, marco ou objeto de referência) e um termo de relação espacial que descreve a configuração geométrica entre os dois objetos (CLARK, 1973). A compreensão dos termos da relação espacial é importante porque reflete a maneira como o ambiente observado é compreendido e descrito pelos usuários por meio da linguagem natural. Diariamente,

as pessoas descrevem cenários e/ou localizações e os termos de relação espacial caracterizam como os elementos dos ambientes se conectam entre si (FERREIRA; DELAZARI, 2019a).

E ainda, vale ressaltar a análise de como as relações espaciais se comportam a fim de formalizar a semântica das expressões locativas (BENNETT; AGARWAL, 2007).

2.2.3 Estudos interlinguísticos de descrições de localização

As comparações da forma e das características das descrições de localização entre os idiomas não foram muito exploradas para muitos idiomas. Uma série de estudos interlinguísticos concentrou-se em preposições espaciais específicas, incluindo situações de contato ou suporte em inglês, japonês e coreano (MUNNICH; LANDAU; DOSHER, 2001) e situações com ou sem contato em superfície física, em inglês e coreano (MCDONOUGH; CHOI; MANDLER, 2003). Além disso, verbos em situações espaciais simuladas têm sido estudados em cenas espaciais estáticas, por exemplo, envolvendo a posição axial com os verbos *standing* e *lying*, em alemão e inglês (BOSSE; PAPAFRAGOU, 2010); eventos usando os verbos *putting* e *taking* em árabe, chinês e inglês (REGIER; ZHENG, 2007) e em suíço e alemão (BERTHELE, 2015); e eventos causais como *the player kicked the ball into the goal* em inglês e grego (BUNGER et al., 2016) (p.1018). Outros investigaram verbos usados para fornecer ou pedir informações de orientação. Por exemplo, Li e Huang (2020), analisou, categorizou e comparou as informações de orientação em direções verbais em inglês e mandarim, e concluiu que os falantes de inglês fornecem mais descrições espaciais do que os falantes de mandarim. Bowerman (1996, p. 393-395) encontrou diferenças em termos de contato e suporte em superfícies verticais e horizontais em inglês, finlandês e espanhol, com situações específicas sendo agrupadas e mapeadas de diferentes maneiras para termos de relações espaciais. Além disso, o mesmo autor apresentou a classificação dos verbos de ação *hanging* e *hooking* em inglês e coreano.

Princípios gerais da linguagem espacial como posição estática, movimento e quadros de referência também foram estudados (GRIGOROGLOU; PAPAFRAGOU, 2018). Os autores revisaram as semelhanças entre a linguagem espacial e a cognição humana em alguns estudos interlinguísticos. Hoffmann (2019) explorou o

uso do quadro de referência absoluto e como se orientar em três línguas australianas, com todas as três exibindo diferentes maneiras de usar relações espaciais.

As análises interlinguísticas também investigaram a interação entre paisagem, cultura, linguagem de localização espacial e topografia. Por exemplo, Brown (2008) estudou o uso de “cima, baixo e transversalmente” em Tzeltal. Este autor descobriu que existem pelo menos três tipos de determinantes da terminologia da paisagem: saliência perceptiva, recursos interacionais e importância cultural. Palmer et al. (2019) investigou quais línguas são moldadas pelo ambiente físico em que são faladas e quais são moldadas por padrões culturais em línguas indígenas australianas, estudando quadros de referências que as línguas utilizam; características topográficas mencionadas; cultura, estilo de vida e atividades da comunidade e classificação do terreno topográfico.

Variações em como as línguas expressam relações espaciais também foram estudadas a partir da perspectiva da aprendizagem de uma segunda língua (COVENTRY; GUIJARRO-FUENTES; VALDÉS, 2012), com desafios para os alunos colocados pela influência do ambiente, quadros de referência e topologia no uso dos termos de relação espacial, juntamente com polissemia e estudo de um grupo de preposições em diferentes línguas. Preposições como *over*, *above*, *under* e *below* em inglês e *za* em russo são bons exemplos das complexidades enfrentadas pelos alunos (TYLER, 2012). As diferenças na posição vertical, contato, suporte e inclusão são amplamente documentadas em estudos com foco no aprendizado de uma segunda língua. Os exemplos incluem: o estudo linguístico cruzado de 17 idiomas usando posições estáticas em desenhos e destacando a gama de diferentes usos de preposições espaciais comuns (FEIST, 2004), e a ausência de uma palavra de contenção para todos os fins equivalente ao inglês *in* em Tzeltal (Língua maia do México) (BOWERMAN, 1996).

Mesmo dentro das famílias de línguas, as diferenças na linguagem espacial são evidentes. Por exemplo, Berthele et al. (2015) estudaram os eventos com verbos de movimento de cinco línguas germânicas e mostrou que existem padrões diferentes na forma como as pessoas expressam sentenças espaciais. Em outro estudo, Majid, Jordan e Dunn (2015) descobriram que a semântica dos termos de relação espacial varia em 12 línguas germânicas muito mais do que outros domínios

que foram testados (cor e partes do corpo), e que os significados em classes próximas eram mais variáveis do que aqueles em classes afastadas.

Torna-se claro a partir desses estudos que os termos de relações espaciais relacionados a contato, suporte e contenção variam amplamente entre os idiomas, dependendo da situação espacial dos objetos envolvidos. Além disso, as línguas não diferem apenas em como classificam as configurações espaciais, mas como entendem as situações do ambiente espacial ao seu redor (BOWERMAN, 1996).

Os SIG são muito utilizados para consultas de localização ou cruzamento de informações e dependem de modelos de relações espaciais como nos exemplos de Clementini, Sharma e Egenhofer (1994) e Cohn et al (1997); mapeamentos simples com termos topológicos em diferentes linguagens são frequentemente realizados, mas abordagens que tragam a linguagem espacial que as pessoas utilizam diariamente (particularmente termos de relação espacial) precisam ser implementadas para melhorar a usabilidade do SIG para falantes de diferentes línguas. Além disso, a geração automatizada de linguagem espacial autêntica requer um entendimento dos elementos que os falantes dessas línguas selecionam como pontos de referência ao descrever a localização e a maneira como aplicam os termos de relação espacial entre estes pontos. Estão surgindo estudos com termos espaciais em análises interlinguísticas examinando semelhanças e diferenças entre duas ou mais línguas diferentes (BATORÉO, 1998; CAMBRIA; WHITE, 2014; COVENTRY; GUIJARRO-FUENTES; VALDÉS, 2012; HOFFMANN, 2019; MARK; GOULD; NUNES, 1989).

2.3 SEMÂNTICA

Noth (1990) em seu breve histórico explica que a etimologia da semântica é derivada do verbo grego *semaínen*. O precursor desta terminologia foi Michel Jules Alfred Bréal que no século XIX atribuiu o nome de semântica para o estudo das leis que governam o significado. O conceito de significado pode variar entre o realismo e o relativismo, onde o realismo explica que a língua se superpõe como uma nomenclatura a um mundo no qual as coisas existem objetivamente, e o relativismo mostra que é a estrutura da língua que determina a capacidade humana de perceber o mundo (ALMEIDA; SOUZA, 2011). Portanto, há duas correntes de estudo em semântica: a realista e a cognitiva. A abordagem realista defende que o significado

de uma expressão se encontra no mundo, enquanto, na cognitiva, o significado de uma expressão coincide com entidades puramente mentais (MARCH CASTAÑEDA, 2017).

Segundo Pinker e Ravagnani (2008) a semântica trata da relação das palavras com os pensamentos, portanto é correto afirmar que o significado de uma palavra depende da cognição humana que pode variar de pessoa para pessoa. Isto, porque na maioria das vezes não se aprende o significado de uma palavra pesquisando-se sobre ela, e sim, ouvindo a palavra dentro de um contexto em que outra pessoa pode definir seu significado. As noções de espaço, tempo, causação, posse e objetivos parecem formar a linguagem cognitivamente. O autor também apresenta a semântica conceitual, a qual mostra que os significados das palavras são representados na mente como montagens de conceitos básicos, em uma linguagem do pensamento. Existem três maneiras de explicar a semântica conceitual: nativismo extremo (ênfase à organização mental inata), pragmática radical (a mente não contém representações fixas dos significados das palavras, depende do contexto) e o determinismo linguístico (a língua nativa é a língua do pensamento).

A semântica é um campo muito complexo, existem diferentes abordagens para seu estudo, e também existem sobreposições entre os estudos da semântica, como por exemplo: filosofia, cognição, linguística (PINKER; RAVAGNANI, 2008), WEB semântica (ALMEIDA; SOUZA, 2011), Inteligência Artificial com o PLN (HERNÁNDEZ-PASTOR; PERIÑÁN-PASCUAL; VALENCIA, 2016) e o geoprocessamento. Parece ser evidente que a construção do significado envolve modelos mentais das perspectivas sobre a realidade. E que o funcionamento das configurações espaço-temporais analisadas e as respectivas utilizações no Português Europeu permitem inferir aspectos muito significativos sobre o funcionamento semântico não apenas para o Português Europeu, mas para todas as línguas naturais. O modelo mental dinâmico que suporta a concessão da escrita em Português (processo temporal entre antes e depois) é diferente do modelo mental de muitas outras línguas (processo espacial). Outro aspecto importante é que a tradução de palavras não deveria ser somente a substituição de palavras, mas sim, as equivalências entre modelos cognitivos que retratem um estado de coisas ou uma realidade percebida (TEIXEIRA, 2013).

O PLN surgiu de uma série de técnicas computacionais para a análise automática e para a representação da linguagem que as pessoas utilizam no dia a dia. No início a análise de uma sentença poderia levar até sete minutos (na época dos cartões perfurados) e atualmente milhões de páginas na *WEB* são processadas em segundos. A evolução da PLN ocorreu e ocorre com sobreposições de estudos e testes em: sintaxe, pragmática e semântica. A sintaxe especifica a maneira como os grupos de símbolos devem ser organizados, de modo que fiquem em formato adequado; a semântica especifica o que as expressões significam e a pragmática especifica como as informações contextuais podem ser aproveitadas para fornecer melhores correlações entre diferentes semânticas, o que é essencial para tarefas de desambiguação do sentido das palavras (CAMBRIA; WHITE, 2014).

A semântica é frequentemente comparada com sintaxe e pragmática. Porém, a sintaxe é o estudo de regras e relações entre palavras e outras unidades em uma linguagem, que determinam o seu arranjo e combinação numa sentença. Alguns linguistas consideram que a sintaxe e a semântica de linguagens naturais estão estreitamente inter-relacionadas. Enquanto a pragmática é o estudo das regras que determinam como diferentes contextos afetam a comunicação do significado (MARCH CASTAÑEDA, 2017).

A combinação da semântica com categorias lexicais torna a representação espacial difícil de ser tratada computacionalmente, e um exemplo disto, são as preposições que representam uma categoria gramatical de uso frequente nos idiomas mundiais. Estudos têm contribuído com progressos no uso de expressões espaciais como as preposições, no que diz respeito ao PLN. Preposições são consideradas uma classe lexical difícil de ser gerenciada e muitos trabalhos fazem abordagens específicas, por exemplo, as tentativas de explicar o modo como os humanos categorizam as preposições espacialmente, do ponto de vista cognitivo-gramatical, em diferentes línguas (ALMEIDA; SOUZA, 2011; HERNÁNDEZ-PASTOR; PERIÑÁN-PASCUAL; VALENCIA, 2016).

Kremer e Baroni (2011) explicam que a semântica desempenha um papel fundamental nos estudos que investigam a representação mental e o processamento dos significados das palavras, especialmente em conceitos e suas categorizações. Normalmente, pesquisadores que trabalham nesta área coletam características semânticas através de uma abordagem experimental em que os participantes são apresentados a um conjunto de conceitos e respondem características que eles

acham que melhor descreveria cada um dos conceitos. Isto permite que os pesquisadores testem teorias sobre memória semântica, para construir estímulos para experimentos posteriores e para modelar o comportamento humano em modelos de simulação computacional. As características dadas por estes tipos de testes são dadas por representações semânticas que acontecem no cérebro humano, podem ser verbalizadas (o que pode trazer uma limitação) e estão sujeitas a variações. As normas semânticas e suas representações conceituais servem como base para uma série de fenômenos inclusive para preparações de similaridades semânticas.

Stock e Pullar (1999) explicam que a semântica que os indivíduos têm para uma determinada entidade do mundo real varia dependendo dos conceitos ou categorias que eles usam para classificar as entidades que encontram na vida cotidiana. Um dos problemas mais difíceis de ser resolvido pela semântica é que as pessoas veem o mundo de maneiras diferentes baseando-se em sua educação e na sua experiência. Diferenças culturais, linguísticas e até de profissão fazem com que as pessoas descrevam o mundo de maneiras diferentes, por exemplo, um ecologista e um engenheiro definiriam um rio de uma maneira completamente diferente. Casos como este mostram que existem e são esperadas diferentes explicações semânticas (STOCK, 2008). A FIGURA 3 ilustra como se pode pensar sobre diferentes corpos d'água e o que caracteriza cada um deles na visão individual das pessoas.

FIGURA 3 - COGNIÇÃO SOBRE DIFERENTES CORPOS D'ÁGUA



FONTE: A autora (2021).

Pode-se dizer que o conjunto de categorias espaciais é limitado e as características espaciais são poucas, por isto suas combinações são restritas. As características das categorias espaciais estão associadas a algumas classes de palavras, tipicamente preposições, verbos, adposições e advérbios. Informações relativas à localização ou movimento de um objeto ou ser vivo podem ser codificadas por um ou mais termos linguísticos dentro de uma frase. Restrições combinatórias e estruturais produzem os esquemas abstratos da localização do espaço e navegação, que pertencem às representações cognitivas e às estruturas linguísticas ao mesmo tempo. Uma questão importante diz respeito ao papel das estruturas linguísticas como fonte de informação semântica, capaz de complementar, ou às vezes até de substituir a experiência sensorial. A linguagem enriquece nossas representações sensoriais, aumentando o poder de nossas estruturas representacionais, embora a linguagem espacial dependa de nossa experiência pré-linguística. Em geral, pode-se pensar que somente o contexto linguístico de palavras tem um papel fundamental na semântica, afirmando que a linguística é rica o suficiente para fornecer uma base semântica para as palavras, como por exemplo, os cegos que frequentemente usam estruturas linguísticas para derivar representações semânticas de termos para os quais eles não podem ter estímulos sensoriais adequados (LANDAU; JACKENDOFF, 1993; MAROTTA, 2010).

Kordjamshidi et al. (2013) explicam que além das características espaciais que estão diretamente disponíveis a partir da linguística, a análise da semântica formal pode fornecer suporte adicional para o mapeamento do idioma para representações qualitativas do espaço. Desta forma, a dependência lexical pode ser reduzida, e categorizações semânticas podem ser usadas como um recurso espacial adicional no processo de mapeamento.

Nos últimos anos, técnicas computacionais para a determinação de características semânticas lexicais foram propostas e avaliadas, incluindo: sentido, registro, domínios específicos, restrições pragmáticas de uso, semântica de marcação e orientação, bem como ligações identificadas automaticamente entre palavras (relacionamento semântico, sinonímia, antonímia e meronímia). Estudos baseados nestes pontos permitem a construção ou ampliação de léxicos, e a atribuição de rótulos semânticos a palavras e frases em textos. A semântica no âmbito da tecnologia da informação, como a *WEB* semântica promove melhorias nos processos de representação e recuperação da informação na *WEB*, a qual, desde

1990, é caracterizada pelo uso de linguagens de marcação e por mecanismos de busca baseados em algoritmos com orientação à sintaxe (HATZIVASSILOGLOU; WIEBE, 2000). Estes algoritmos que recuperam textos *online* da *WEB* são capazes de dividir as sentenças em partes, verificar a ortografia e contar as palavras, mas quando se trata de interpretação e extração de informações significativas, são muito limitados. Portanto, um PLN deve trabalhar com (CAMBRIA; WHITE, 2014):

- A criação e propagação de ligações dinâmicas,
- Manipulação de estruturas,
- Aquisição e acesso ao léxico e semântico da língua em questão,
- Controle dos módulos de aprendizado e processamento e roteamento das informações entre estes dois módulos,
- Construção do nível básico da linguagem (como por exemplo: objetos e ações) através da percepção e experiências motoras, e por fim,
- A representação dos conceitos abstratos.

Como ainda não se tem todos estes itens para serem utilizados na *WEB*, basicamente, duas características da *Internet* dificultam o acesso à informação útil, específica e relevante: o volume e a falta de definição semântica precisa e interpretável por programas e sistemas, para as informações disponibilizadas nas páginas. A falta de mecanismos capazes de captar a semântica do conteúdo das páginas da *WEB* criou uma demanda de serviços que se ajusta adequadamente à classe de serviços estudada em Inteligência Artificial. Uma das propostas de solução para a problemática citada seria dotar a internet de inteligência, fazendo com que as páginas possuam uma semântica clara e definida, e que agentes possam raciocinar sobre esta semântica. Essa ideia deu origem à *WEB Semântica*, onde as ontologias têm um papel fundamental (DE FREITAS, 2003).

Ferreira, Claro e Lopes (2011), explicam que descrições semânticas vêm sendo incorporadas em serviços *WEB* e também em ambientes integrados de desenvolvimento (IDE) com o intuito de facilitar a utilização por desenvolvedores. Dentre as linguagens semânticas, a OWL-S (*Ontology Web Language for Services*) permite descobrir e selecionar serviços *WEB* automaticamente, pois adiciona características não ambíguas e interpretáveis computacionalmente. Os serviços

WEB que utilizam características semânticas na sua descrição são denominados Serviços *WEB* Semânticos (SWS).

Vale a pena ressaltar as colocações de De Freitas (2003) sobre a não utilização da semântica para capturar o conhecimento contido nas páginas da *internet*. Pois, devido a este fato, os mecanismos de busca caracterizam-se por uma alta cobertura, porém com significativa falta de precisão, muitas vezes entregando ao usuário uma grande quantidade de endereços de páginas inúteis ou irrelevantes. Os mecanismos de busca utilizam algoritmos matemáticos para atribuir relevância às páginas, e não conseguem aplicar a semântica na busca porque possuem capacidade de representar as páginas com análises baseadas apenas no nível léxico. Eles indexam as páginas da *Internet* por palavras-chave, e aplicam métodos e estruturas de dados para recuperá-las rapidamente, devolvendo ao usuário uma lista de endereços de páginas que contém as palavras solicitadas, ordenadas por frequência e não por processos de semântica.

O processo de pesquisa para dados geoespaciais na *WEB* também é baseado em palavras-chaves. Porém os mecanismos de pesquisa têm dificuldades com as pesquisas de dados espaciais ao procurar palavras-chaves, porque não consideram a semântica das estruturas de linguagem natural (TOMKO; WINTER, 2005). Aplicar a semântica nestes dados de geoprocessamento permitiria e facilitaria a descoberta de informações que estes dados contêm. A aplicação de conceitos em dados geoespaciais facilitaria a busca de dados para compor bases cartográficas de acordo com o que os usuários estejam procurando. Estas aplicações contribuiriam com a demanda da criação de bases cartográficas. Yue et al. (2011), discutem a criação de um catálogo de serviços geoespaciais baseado em semântica e para isto, apontam alguns itens importantes: a) representação da semântica para os dados geoespaciais usando a abordagem da ontologia; b) incorporação da semântica em diversos banco de dados existentes aumentando assim os serviços catalogados; c) processamento de consultas do catálogo usando semântica; d) aplicação de conceitos nas informações geoespaciais que possibilitem a integração dos dados e serviços existentes no catálogo. Os autores explicam que a resolução dessas questões pode ser de grande valor para lidar com a heterogeneidade semântica no suporte ao acesso a dados geoespaciais, processamento de informações e descoberta de conhecimento, contribuindo para a evolução da infraestrutura na

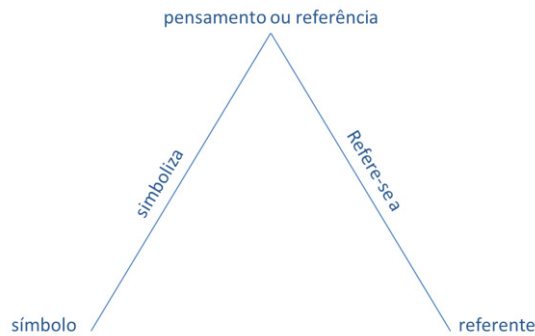
WEB. Estudos neste porte mostram a importância da semântica para dados geoespaciais.

Um exemplo de aplicação para resolver problemas relacionados às dificuldades que usuários de dados geoespaciais da Infraestrutura Nacional de Dados Espaciais do Brasil (INDE) enfrentam, quando necessitam encontrar dados potencialmente úteis, foi o estabelecimento de uma estrutura de integração semântica através do uso de ontologias. Ou seja, problemas em avaliar a adequação dos dados existentes e em como integrá-los a uma aplicação de interesse, se resolveriam se os dados fossem estruturados de forma inteligente e que pudessem ser interpretados através de seus significados, com isto, entra a questão da interoperabilidade semântica, que aprimora a eficácia dos sistemas de busca, propõe avaliação e integração da geoinformação oferecida por fontes diferentes, como é o caso da INDE (MARCH CASTAÑEDA, 2017).

A semântica estabelece uma ponte entre a linguagem e o pensamento; e sugere o significado atrelado aos objetos do mundo real (*referentes*) e aos conceitos (*referências*). Portanto, existe uma diferença fundamental entre conceitos como “coisas” deste mundo e conceitos como “representações” (produzidas pela mente) de coisas, que são frequentemente (de modo intencional ou acidental) parciais e dependentes de contexto (MARCH CASTAÑEDA, 2017).

A representação da semântica pode ser dada pelo triângulo do significado que tem nos seus cantos o símbolo, o pensamento ou referência e o referente (FIGURA 4). A relação entre o símbolo e o pensamento é causada (simboliza) pelo objetivo para o qual se está fazendo uma referência, ou seja, quando se ouve o que é dito, os símbolos levam o ouvinte a realizar um ato de referência e que deverá ser semelhante ao que o falante quer expressar. Portanto, a relação entre o pensamento e o referente é direta e o entre o símbolo e referente não há uma conexão direta (OGDEN; RICHARDS, 1923). O triângulo semiótico passou por uma longa série de interpretações e modificações, em diferentes teorias e autores (MARCH CASTAÑEDA, 2017).

FIGURA 4 – TRIÂNGULO DO SIGNIFICADO



FONTE: Adaptado de Ogden e Richards (1923, P.11).

A aplicação de conceitos ou conceituação se refere a representações que dão sentido ao ambiente onde os elementos e/ou pessoas estão localizados. Portanto, quando existir um alinhamento final de todas as conceituações espaciais haverá uma convergência para a formação de um modelo global para que toda a humanidade possa utilizar, como é o caso da matemática que é de consenso universal. No entanto, conceituações sobre um domínio específico podem depender do contexto social e/ou geográfico. No caso das relações espaciais, até os termos mais utilizados, como é o caso das preposições, variam de idioma para idioma, tendo que ser estudados em suas línguas nativas. Um exemplo disto é o termo “em” (in) que em inglês significa “contido” e em coreano significa “fricção”. Estas discrepâncias nas linguagens mostram que parece improvável que ocorra a formação global dos termos de relações espaciais, porém permitem analisar que é possível ter diferentes conceituações do mesmo domínio sem a necessidade de correção de uma em relação à outra. Portanto, a conceituação de dados geoespaciais não serve apenas para comparar os termos que são utilizados em um determinado contexto, cultura ou idioma e sim que contribua com noções geográficas básicas, como por exemplo, encontrar a localização de elementos, pessoas ou lugares (TANASESCU, 2007).

Os sistemas conceituais das línguas sofrem transformações ao longo do tempo, portanto a estrutura destes sistemas também pode sofrer alterações, sejam elas fonológica, morfológica, lexical, sintática, semântica ou discursiva. O número de possibilidades de codificação espacial nas mais diversas línguas mostra como é difícil explicá-las por uma única teoria, pois envolve fatores intra e extralinguísticos (KEWITZ, 2011). Além do mais numa visão generalista, a realidade física (veículos,

peças, cidades) e até a realidade não física (densidade populacional, renda média, volume de tráfego) têm referências geoespaciais (MARCH CASTAÑEDA, 2017).

O conceito é a unidade básica da semântica (RHEE, 2004) por isto pode ser descrito por uma definição que se refere a propriedades do conceito ou por uma coleção de indivíduos representativos do conceito. Relações espaciais geralmente associam dois conceitos ou mais, de modo que é improvável que se refiram à posição absoluta (localização) de instâncias dos geoconceitos. Relações topológicas são mais comuns para serem consideradas (MARCH CASTAÑEDA, 2017), porém não são suficientes para as representações dadas pelas linguagens naturais.

A semântica é útil na teoria e na prática para lidar com os significados das palavras porque permite comparações no que diz respeito a semelhanças e diferenças. Na investigação das 20 preposições utilizadas com mais frequência na língua inglesa, Rhee (2004) descobriu que o significado de “locação” foi o mais utilizado, seguido pelo significado de “movimento”.

2.4 SIMILARIDADE SEMÂNTICA

Existem vários tipos de relações semânticas: sinônimos, metonímias, antônimos, associações, entre outros. Os sistemas de PLN precisam de relações semânticas por causa da lacuna entre a superfície lexical do texto e seu significado, pois, dois documentos de texto podem descrever a mesma entidade com termos diferentes. As relações semânticas de sinônimos, hiperônimos e hipônimos, mostraram-se úteis para aplicações em PLN que incluem consultas, perguntas e respostas, desambiguação de palavras e similaridade (PANCHENKO, 2013).

Seno e Nunes (2009) explicam que a similaridade entre palavras é obtida a partir de um conjunto de sinônimos, enquanto a similaridade entre multipalavras é determinada com o uso de um léxico de paráfrases. A construção de um léxico representativo de paráfrases requer um grande volume de dados, o que é um recurso praticamente inexistente para a maioria das línguas. Podem-se utilizar as paráfrases para comparação de sentenças e seu tratamento automático ainda é um desafio. Uma aplicação do uso de regras de paráfrases é em análises de *corpus* com o conhecimento lexical, sintático e semântico (isto é, relações de sinonímia) que

possibilitam a identificação de palavras e multipalavras que conduzem informações semanticamente similares em sentenças do português.

A fusão de sentenças é uma área de pesquisa em PLN e é motivada por aplicações práticas tais como a Tradução Automática, a Sumarização Automática e os sistemas de Perguntas e Respostas. Na sumarização de documentos, por exemplo, o processo de fusão de informações comuns é de grande relevância para eliminar a redundância de informações nos sumários. A fusão de sentenças é comumente dividida em três etapas: identificação de informações comuns, fusão de informações e linearização. A primeira etapa consiste em reconhecer informações semanticamente similares (por exemplo, paráfrases e sinônimos) que se repetem nas sentenças. A segunda etapa consiste em escolher os itens lexicais que irão compor a nova sentença e determinar o modo como eles serão combinados na sentença. A última etapa, por sua vez, consiste em realizar em língua natural a sentença obtida a partir da etapa anterior e envolve, portanto, aspectos gramaticais da sentença (SENO; NUNES, 2009).

Medições de similaridade semântica em textos também tem sido assunto de pesquisas em PLN, recuperação de informações e inteligência artificial com tarefas que incluem: reconhecimento de paráfrase, buscas no *Twitter*, recuperação de imagens por legendas, reformulação de consulta, tradução e avaliação automática e esquemas de correspondência. O crescimento exponencial das informações descritivas na *WEB* criou a necessidade de desenvolver novas ferramentas de PLN para processar e gerenciar automaticamente o conteúdo disponível na Internet. Além dos gêneros textuais de noticiários e pesquisas, existem os blogs, fóruns e as redes sociais que utilizam contextos descritos em linguagem natural. Estes exemplos mostram a dificuldade de desenvolver ferramentas de PLN, devido à alta variabilidade semântica da linguagem (HAN et al., 2013).

Moldovan et al. (2015) explicam que no estudo da memória semântica leva-se em consideração a questão de como o significado é representado na mente/cérebro do ser humano. As teorias da memória semântica podem ser divididas em duas classes: o significado é representado de maneira holística e o significado da palavra é representado por um modelo de distribuição. Modelos distribuídos de memória semântica assumem que o significado da palavra é representado por um conjunto de nós, cada um representando uma característica (ou micro característica) de um conceito. Em outras palavras, o significado é

decomposto em muitas características diferentes, e cada uma delas podem ser parte do significado de múltiplos conceitos. As representações holísticas são de nós únicos em uma rede semântica. Embora as visões holísticas e os modelos distribuídos tenham maneiras diferentes de conceituar um significado, ambas as abordagens assumem que os componentes críticos dos conceitos têm propriedades ou características semelhantes e que dependem dos aspectos que eles têm em comum. De fato, a similaridade semântica entre os conceitos lexicais tem se tornado uma das principais questões teóricas do estudo de representação de significado.

Stock e Pullar (1999) explicam que para que um banco de dados seja integrado a outros bancos de dados com sucesso, seus elementos devem ser similares semanticamente. No entanto, pode haver diferenças entre os conceitos que os usuários possuem para a mesma entidade do mundo real. Segundo os autores, os benefícios da partilha de dados espaciais são discutidos desde a década de 1980, porém as tentativas de atingir o objetivo foram frustrantes pela heterogeneidade dos dados (esquemática, sintática, semântica). A heterogeneidade semântica refere-se a diferenças na definição de conceitos e as regras que são usadas para determinar se uma entidade do mundo real é um exemplo de um conceito.

Stock e Pullar (1999) demonstram que existem dois grupos de métodos para determinar a similaridade semântica dos elementos de um banco de dados. Os métodos no primeiro grupo aproveitam a representação de bases de dados de várias maneiras, comparando representações para identificar elementos semanticamente equivalentes. Os métodos no segundo grupo tentam identificar elementos semanticamente equivalentes interrogando definições ou relações entre termos. Esses métodos são limitados em sua capacidade de identificar semelhanças e diferenças, pois geralmente assumem que uma linguagem comum é usada (ou seja, um termo tem o mesmo significado em todas as bases de dados). Esta suposição é muitas vezes inválida no contexto de dados espaciais, porque o mesmo termo pode ser usado por diferentes usuários para significar coisas diferentes. Da mesma forma, diferentes termos podem ser usados por diferentes usuários para significar a mesma coisa.

De acordo com Stock (2008) os serviços *WEB* geoespaciais estão cada vez mais disponíveis e acessíveis através da IDE. No entanto, muitas vezes é difícil para os usuários encontrarem e interpretarem Serviços *WEB* adequados para atingir os

seus objetivos sem intervenção manual significativa. As ontologias de serviços da *WEB* têm tentado representar uma semântica na esperança de que seja possível para o usuário alcançar seus objetivos automaticamente. No entanto, é necessário um grande esforço na compreensão de ontologias para que a semântica dos serviços *WEB* seja especificada de acordo com os requisitos do usuário.

Ferreira, Claro e Lopes (2011), explicam que cada parâmetro utilizado nos serviços *WEB* é mapeado em uma classe de uma ontologia de domínio. É importante ressaltar que a utilização de ontologias é essencial para a fase de substituição por termos semanticamente similares, pois é por meio delas que serão efetuadas as inferências necessárias para identificar a correlação entre as classes.

Moldovan et al. (2015) explicam que num sentido mais amplo, palavras principais e os alvos podem ser considerados semanticamente relacionados se forem membros da mesma categoria, se eles forem sinônimos, se tiverem uma relação funcional, ou se eles têm uma relação parte-todo. Um tipo diferente de relacionamento que tem sido amplamente examinado na literatura é o *semantic-priming*, que são as associações entre uma palavra principal e uma palavra alvo (por exemplo, rato - queijo). Enquanto a relação semântica reflete a semelhança de significado entre dois conceitos, as relações associativas entre pares de palavras baseiam-se principalmente na co-ocorrência e são pensadas para refletir o uso da palavra em vez da sobreposição de significados.

Os termos espaciais são altamente polissêmicos, tem mais de um significado, por exemplo, a preposição *on* da língua inglesa, não determina somente um relacionamento espacial, mas vários: *a book on a table* (um livro sobre a mesa), *a picture on a wall* (um quadro em uma parede) ou *an apple on a branch* (uma maçã num galho). Já a língua holandesa, apesar de sua proximidade com a língua inglesa, trata os três casos anteriores com preposições diferentes: *op*, *aan* e *om* respectivamente. Não que a língua inglesa sempre agregue as coisas e as outras línguas as separem, isto não é regra. Muitas línguas não distinguem *on* de *over*, usando o mesmo termo para sobre com adjacência e sem adjacência, ou não distinguem *in* (em/dentro de) de *under* (sob/debaixo). As línguas tendem a possuir termos para contato, alinhamento vertical, proximidade e para estar contido ou não, como se existisse um alfabeto cognitivo mais básico de relações espaciais que as preposições de cada língua. Então, quando os idiomas juntam várias relações espaciais na mesma preposição, aderem a uma noção universal sobre quais

relações são mais semelhantes entre si. Retornando ao exemplo *on* da língua inglesa, o *on* junta um livro sobre uma mesa (alinhamento vertical e contato) com um quadro na parede (ligação), presumidamente porque ambos envolvem uma força que mantém uma coisa em contato com a outra. A língua espanhola junta todos os três no *en*. Nenhuma língua junta o alinhamento vertical com o “estar contido ou não” ao mesmo tempo em que se exclui a ligação, ou o estar por cima e o estar perto excluindo o estar sobre, porque essas junções não fazem sentido em termos cognitivos (PINKER; RAVAGNANI, 2008).

De acordo com Hatzivassiloglou et al. (1999) apud (Seno; Nunes (2009), duas sentenças são semanticamente similares se elas se referem a um mesmo objeto ou evento. E ainda, se o objeto realiza a mesma ação em ambas as sentenças, ou é sujeito da mesma descrição. Os autores citam como exemplo, as sentenças (a), (b) e (c):

- a) Uma bomba caseira foi atirada contra a sede do Ministério Público (MP).
- b) Uma bomba caseira foi jogada contra o prédio do Ministério Público, na capital do estado.
- c) Uma bomba caseira atingiu o prédio da Secretaria de Estado da Fazenda, localizado na Avenida Rangel Pestana, ao lado do Poupatempo Sé.

Apesar de todas as sentenças se referirem à explosão de uma bomba caseira, as sentenças (a) e (b) focam na explosão ocorrida no Ministério Público, enquanto (c) se refere à explosão ocorrida na Secretaria de Estado da Fazenda. Nesse caso, somente (a) e (b) são consideradas similares.

De acordo com Hoey (1991) apud Seno; Nunes (2009) duas sequências distintas de palavras são ditas paráfrases se uma delas puder ser substituída pela outra, em um dado contexto, sem alterar significativamente o sentido do texto.

Um exemplo de aplicação de similaridade semântica de relações espaciais foi o estudo de um sistema de diálogo para gerar instruções de rota *indoor* em alemão. A tarefa desafiadora do estudo foi de elaborar um conjunto compacto de instruções precisas e compreensíveis. O sistema tem quatro módulos principais: compreensão da linguagem natural, gerenciamento de diálogo, geração de instruções de rota e geração da linguagem natural. Para este sistema, uma parte do

estudo da língua alemã compreendeu a similaridade semântica de sentenças. Por exemplo, a simples ação de girar para a esquerda compreendeu um subconjunto de cinco frases possíveis:

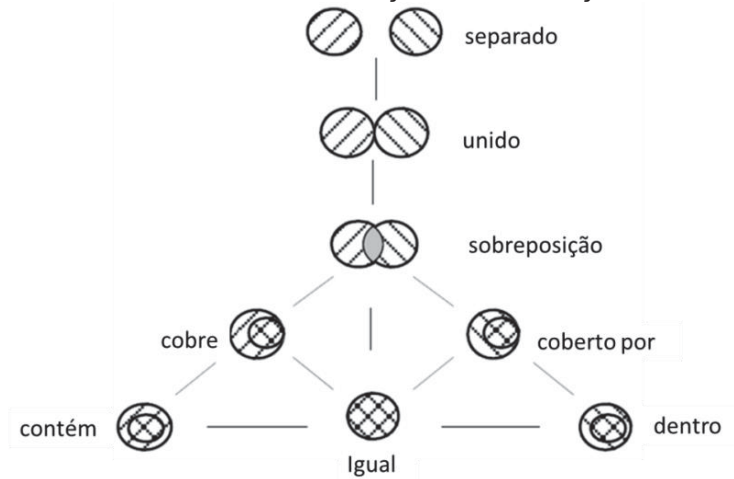
- 1) Vire à esquerda.
- 2) Vire para a esquerda.
- 3) Você está virando à esquerda.
- 4) Esquerda.
- 5) Vá para a esquerda.

As frases variam do modo imperativo para o modo declarativo e na flexão do verbo. Estas variações existentes em forma de similaridades semânticas foram capturadas e estruturadas para receberem um único formalismo (CUAYÁHUITL et al., 2010).

O artigo de Bruns e Egenhofer (1996) enfoca quais medidas de similaridade que são apropriadas para a recuperação de configurações espaciais semelhantes expressas em uma linguagem de consulta espacial para cenas em um BDG. Uma cena espacial é um conjunto de objetos geográficos juntamente com suas relações espaciais (topológicas, de distância e de direção) e, opcionalmente, outros tipos de características espaciais como forma, proporção, tamanhos relativos (áreas e comprimentos), ou atributos que especificam a semântica dos objetos. Avaliar a semelhança entre diferentes cenas envolve comparar tuplas ou conjuntos de tuplas no BDG. Para isto, os autores impuseram uma ordem de relações espaciais para que acontecesse uma mudança gradual dos objetos até que a relação espacial entre eles fosse alterada. Desta maneira, as cenas espaciais que necessitaram de pouca mudança para alterar as relações topológicas, de distância e de direção de uma para a outra foram consideradas mais semelhantes às cenas que exigiram mais mudanças. A aplicação de um processo gradual e espacialmente consistente pode transformar uma relação em qualquer um de seus vizinhos conceituais. A FIGURA 5 mostra um exemplo de uma imposição gradual dos conceitos de vizinhança para relações topológicas. A FIGURA 6 mostra um possível esquema de relação de distância com base no aumento das distâncias de buffer, zero, estar muito perto, perto e longe. A FIGURA 7 mostra os esquemas de relação de direção que são altamente sensíveis à orientação dos objetos em uma cena. Os autores concluem o

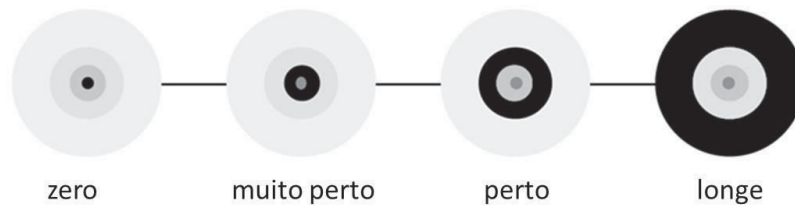
artigo dizendo que existe a necessidade de: a) testar se o modelo escolhido corresponderia às relações espaciais em linguagem natural, b) a classificação de cenas por níveis de similaridade e c) complementar os modelos com a semântica dos objetos envolvidos a fim de fornecer um raciocínio espacial inteligente para ser utilizado computacionalmente.

FIGURA 5 – CONCEITOS DE VIZINHANÇA PARA RELAÇÕES TOPOLÓGICAS



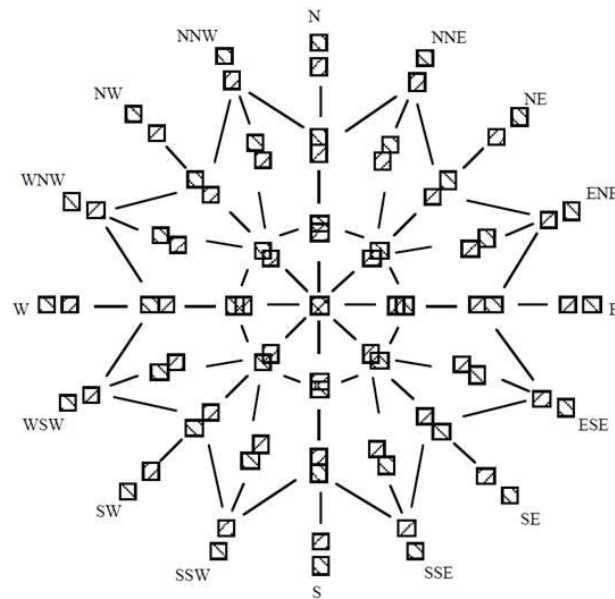
FONTE: Adaptado de Bruns e Egenhofer (1996).

FIGURA 6 – RELAÇÃO DE DISTÂNCIA PARA RELAÇÕES TOPOLÓGICAS



FONTE: Adaptado de Bruns e Egenhofer (1996).

FIGURA 7 – RELAÇÃO DE DIREÇÃO PARA RELAÇÕES TOPOLÓGICAS



FONTE: Adaptado de Bruns e Egenhofer (1996).

O estudo de Bennett e Agarwal (2007) analisa termos semanticamente similares em linguagem natural da palavra “lugar”. Os autores explicam que esta palavra se manifesta de várias maneiras e por isto foi categorizada em 4 classes de acordo com a forma gramatical que foi utilizada. Em substantivos, que podem se referir a classes de objetos físicos (cidade, sala, floresta, prédio) ou a tipos de entidades espaciais (área, vizinhança) ou ainda a substantivos mais abstratos como “local”. Em propriedades locativas que são expressões que caracterizam a localização dos objetos, geralmente contêm uma preposição referente à relação espacial e a referência ao objeto a ser localizado (Em Londres, na colina, embaixo da mesa, entre a igreja e o cinema). Em nomes de lugares (Paris, Inglaterra) e em descrições definidas que funcionam como expressões nominais mais complexas dos lugares (a biblioteca, o galpão na fazenda). E, interessante citar que neste estudo os autores fizeram a classificação explicando as classes de substantivos, nomes de lugares e descrições definidas, deixando de fora as expressões locativas devido à complexidade e diversidade dos termos utilizados em linguagem natural. Isto comprova que as línguas naturais contêm muitos termos de relações espaciais e as pessoas as empregam de acordo com o contexto em que estão inseridas.

Um sistema verdadeiramente inteligente precisa de conhecimento físico de como os objetos se comportam, conhecimento social de como as pessoas

interagem, conhecimento sensorial de como as coisas se parecem e conhecimento psicológico sobre a maneira como as pessoas pensam e assim por diante. Ter um banco de dados de milhões de fatos do senso comum, no entanto, não é suficiente para a compreensão da linguagem natural computacional, necessita-se ensinar aos sistemas de PNL como lidar com esse conhecimento, mas também interpretar emoções e nuances culturais (CAMBRIA; WHITE, 2014), assim a descoberta de termos semanticamente similares torna-se um passo nesta conquista.

Embora a igualdade e a desigualdade sejam fáceis de serem definidas computacionalmente, a similaridade é uma medida vaga e relativamente indefinida para comparar entidades, o que torna sua formalidade complicada. Entretanto, a similaridade desempenha um papel importante na cognição humana: o senso de similaridade é a base da capacidade humana de classificar entidades semelhantes, de raciocinar em situações semelhantes e também de aprendizado, por isto sua importância é fundamental nos estudos relacionados às relações espaciais em linguagem natural. Medir as similaridades semânticas permite que uma máquina compare instruções dadas por seres humanos, procure instruções semelhantes no passado e reutilize para novas procuras. As comparações de similaridade são importantes para simular os processos de raciocínio humano e oferecer a imprecisão e a flexibilidade que o raciocínio lógico não tem. A capacidade de julgar a similaridade é necessária para reagir adequadamente a novas situações, comparando-as com as experiências aprendidas no passado (SCHWERING, 2007). E assim, neste sentido, é fundamental compreender o processo de similaridade como um processo cognitivamente muito mais complexo do que a simplicidade da "semelhança" lexical. Desta maneira, as análises linguísticas devem entender as línguas naturais como processos que estão na dimensão cognitiva humana, na interdependência entre o significado, as conceitualizações e as inferências pragmáticas (TEIXEIRA, 2013).

2.5 ONTOLOGIA

A palavra ontologia é de origem grega e significa *ontos* (ser) + *logos* (conhecimento sobre), tendo sido criada entre os séculos XVII e XVIII por filósofos alemães para denominar o ramo da filosofia que trata da natureza e organização do

ser. O termo ontologia pode ter vários significados dependendo da realidade em que esteja sendo aplicado, visto que é usado para descrever domínios e estes podem abordar um determinado assunto muito específico da sua realidade. Nesse sentido, o termo ontologia foi utilizado inicialmente na filosofia, passando, contudo a ser utilizado também em outras ciências, como por exemplo, a partir da década de 90 do século passado, o termo foi adotado pelas áreas da Inteligência Artificial, Engenharia do Conhecimento, Sistemas de Informação e Ciência da Computação (Recuperação de informações na Internet, PLN, Gestão de Conhecimento, *WEB* Semântica, Educação), onde recebeu adaptações em sua definição (SABINO; HEINZLE, 2015; SOUSA et al., 2011).

A ontologia de referente seria uma teoria sobre como o universo se encontra organizado e corresponde ao ponto de vista do filósofo. A ontologia epistemológica ajusta-se aos propósitos da Ciência da Computação, Ciência da Informação e das áreas de pesquisa de Engenharia do Conhecimento e Inteligência Artificial e se define como uma teoria sobre como um indivíduo, grupo, linguagem ou ciência conceitualiza um determinado domínio (MARCH CASTAÑEDA, 2017).

Sousa et al. (2011) explicam que as ontologias são utilizadas em diversas ciências e são formas de armazenar dados sobre um determinado domínio com o objetivo de realizar a padronização dos mesmos, proporcionando um melhor entendimento sobre os conhecimentos armazenados e facilitando o acesso às informações. A utilização de ontologias em várias áreas do conhecimento tem se tornado uma atividade muito útil, pois são capazes de armazenar, compartilhar e reutilizar conhecimentos e estruturas de informação referentes a um determinado domínio, permitindo a origem de análises e consultas relacionadas aos conhecimentos. Contudo, ontologias devem possuir um significado e abrangência muito mais profundos do que as simples hierarquias de conceitos e palavras-chaves empregadas por muitos mecanismos de busca (DE FREITAS, 2003; NOY; MCGUINNESS, 2001). As ontologias são mais que ferramentas de busca e armazenamento, elas possuem funcionalidades que permitem que a máquina possa processar o raciocínio automatizado, através de regras e inferências (SALES; CAMPOS; GOMES, 2008).

Noy e Mcguinness (2001) explicam que no final da década de 1990, o desenvolvimento de ontologias saiu dos Laboratórios de Inteligência Artificial e foi para os desktops de especialistas em domínio. As ontologias na *WEB* variam de

grandes taxonomias categorizando sites (como no *Yahoo!*) para categorizações de produtos para venda e suas características (como na *Amazon.com*). Diversas áreas desenvolvem ontologias padronizadas que especialistas em domínios podem utilizar para compartilhar e anotar informações em seus campos. Uma ontologia define um vocabulário comum para pesquisadores que precisam compartilhar informações sobre um domínio. Inclui definições interpretáveis por máquina de conceitos básicos de um domínio e as relações entre eles. Segundo os autores, algumas razões para o desenvolvimento de uma ontologia são:

- Compartilhar o entendimento comum da estrutura de informação entre pessoas ou agentes de *software*,
- Permitir a reutilização de conhecimento de um domínio,
- Tornar as suposições de um domínio explícitas,
- Separar o conhecimento do domínio do conhecimento operacional,
- Analisar o conhecimento do domínio

Stock (2008) explica que a construção de ontologias para serviços da *WEB* requer um grande esforço manual e conhecimento especializado para o desenvolvimento de tarefas de especificações lógicas e algébricas.

Para desenvolver ontologias existem vários ambientes, metodologias e linguagens que podem ser utilizados dependendo do tamanho e complexidade do domínio a ser modelado (BREITMAN, 2004; NOY; MCGUINNESS, 2001). Um exemplo é a linguagem OWL, lançada pelo *World Wide Web Consortium* (W3C), projetada para atender as necessidades das aplicações para *WEB* semântica. Um exemplo para a linguagem de *frames* usada na definição de ontologias é a *F-Logic*, que permite ainda a integração de objetos procedimentais e a definição de axiomas (sentenças consideradas sempre válidas, e que, na prática, servem aos atributos como restrições complexas e que podem envolver classes, instâncias e outros atributos) (DE FREITAS, 2003). Um exemplo de ambiente usado para modelagem é o *framework Protégé*, serve para construção e edição de sistemas baseados em conhecimento e possui a possibilidade de carregar, editar e salvar ontologias em vários formatos, tais como, *Resource Description Framework* (RDF), *Extensible Markup Language* (XML), OWL, entre outros (SOUSA et al., 2011).

Sabino e Heinzle (2015) explicam que existem múltiplos formalismos e ferramentas para representar computacionalmente o conhecimento. Todos têm como desafio central a construção de dispositivos computacionais capazes de simular, na máquina, algumas características da inteligência humana. Diferente dos dados, que são armazenados em bancos de dados, o conhecimento necessita de estruturas mais complexas para seu armazenamento. Estas estruturas são chamadas de Base de Conhecimento e são criadas fazendo-se uso de formalismos próprios, dentre os quais se destaca a ontologia.

Segundo De Freitas, 2003; Gruber, 1993; Noy e McGuinness, 2001; Sales, Campos e Gomes (2008), a representação formal de um conhecimento é baseada em uma conceituação (objetos, conceitos, subclasses, relações, funções, axiomas e instâncias) que é uma visão abstrata e simplificada do mundo que se deseja representar para algum propósito. Uma ontologia junto com um conjunto de classes constitui uma base de conhecimento. Toda base de conhecimento ou sistema baseado em conhecimento está comprometido com alguma conceituação, explícita ou implicitamente. Na realidade, há uma linha tênue onde a ontologia termina e a base de conhecimento começa. As classes são o foco da maioria das ontologias, pois elas que descrevem os conceitos do domínio. Em termos práticos, o desenvolvimento de uma ontologia inclui: a) definir as classes da ontologia, b) organizar as classes em uma hierarquia taxonômica (subclasse-superclasse), c) definir os atributos e descrever os valores permitidos para esses atributos, d) preencher os valores para os atributos das classes. Normalmente, uma ontologia é organizada em hierarquias de conceitos (ou taxonomias). Existem requisitos do termo ontologia que possibilitam sua aplicação em informática, conforme mostrado abaixo:

- Especificação explícita: no contexto da *WEB* semântica uma ontologia é uma especificação explícita de uma conceituação, é uma explicação sistemática da existência do que se quer representar. É o entendimento das definições, instâncias, relações, restrições e axiomas.
- Formal: Para sistemas baseados em conhecimento, o que “existe” é exatamente aquilo que pode ser representado. Quando o conhecimento de um domínio é representado em um formalismo declarativamente definido, portanto, compreensível para agentes e sistemas.

- **Conceitualização:** trata de um modelo abstrato de uma área de conhecimento ou de um universo limitado de discurso. O conjunto de objetos e as relações descritíveis entre eles são refletidos no vocabulário representacional com o qual um programa baseado em conhecimento representará o conhecimento desejado. Assim, pode-se descrever uma ontologia de um programa definindo um conjunto de termos representacionais. Nessa ontologia, as definições associam os nomes de entidades no universo do discurso (por exemplo, classes, relações, funções ou outros objetos) com textos descrevendo os significados e conceitos, e os axiomas formais que restringem a interpretação e o uso desses termos.
- **Compartilhada:** conhecimento consensual, podendo ser uma terminologia comum da área modelada ou acordada entre os desenvolvedores dos agentes que se comunicam.

De Freitas (2003) e Sousa et al. (2011) explicam que a classificação de uma ontologia é fundamentada de acordo com o tratamento dado à informação sobre a qual a ontologia está sendo aplicada e de acordo com seu grau de genericidade.

- **Ontologias de representação:** definem as primitivas de representação, como frames, axiomas, atributos e outros, de forma declarativa.
- **Ontologias gerais (alto nível):** descrevem conceitos mais gerais, definições abstratas necessárias para a compreensão de aspectos do mundo, como tempo, processos, papéis, espaço, seres, coisas, etc. São a princípio independentes de domínio e podem ser reutilizadas em novas ontologias.
- **Ontologias centrais ou genéricas de domínio:** definem os ramos de estudo de uma área e/ou conceitos mais genéricos e abstratos desta área. Por exemplo, uma ontologia central de direito que inclui conhecimentos normativos, de responsabilidade, reativos, de agências legais, comportamentos permitidos. Esses conceitos e conhecimentos foram agrupados em uma ontologia para que ela servisse de base para

a construção de ontologias de ramos mais específicos do direito, como direito tributário, de família e outros.

- Ontologias de domínio: descrevem o vocabulário relativo a um domínio específico por meio da especialização de conceitos das ontologias de alto nível, como por exemplo, microbiologia, direito tributário.
- Ontologias de tarefas: descrevem o vocabulário relativo a uma tarefa genérica ou atividade por meio da especialização de conceitos das ontologias de alto nível.
- Ontologias de aplicação: procuram solucionar um problema específico de um domínio, como identificar doenças do coração, a partir de uma ontologia de domínio de cardiologia. Normalmente, ela referencia termos de uma ontologia de domínio.

Noy e McGuinness (2001) explicam que não existe um método ou metodologia “correta” para o desenvolvimento de ontologias. Depois de se definir uma versão inicial da ontologia, certamente precisará revisar todo o processo e essa iteração provavelmente continuará durante todo o ciclo de vida da ontologia. Ao longo do desenvolvimento da ontologia, discutem as decisões de modelagem, bem como os prós, contras e implicações de diferentes soluções. Algumas regras fundamentais no desenvolvimento de ontologias que podem ajudar a fazer decisões são:

- 1) Não há uma maneira correta de modelar um domínio - há sempre alternativas viáveis. A melhor solução quase sempre depende da aplicação que se tem em mente e as extensões que se necessitam.
- 2) O desenvolvimento de ontologias é necessariamente um processo iterativo.
- 3) Conceitos na ontologia devem estar próximos a objetos (físicos ou lógicos) e relacionamentos em seu domínio de interesse. Estes são mais prováveis de serem substantivos (objetos) ou verbos (relacionamentos) em frases que descrevem o seu domínio.

Um exemplo de ontologia de características complexas é dado por Varanka e Jerris (2010) que explicam que estas ontologias controlam a conexão de

elementos entre sistemas de classificação, aplicando o raciocínio lógico. Em uma ontologia complexa de topografia, o contexto geográfico pode ser definido como representação de tempo e lugar de uma determinada configuração. Uma maneira comum de representar a escala geográfica é usar a taxonomia de classificação como uma forma de maior ou menor representação com poucas ou mais instâncias. Para uma modelagem mais rica, a escala operacional dependerá das propriedades da relação espacial modeladas na ontologia e na semântica que elas transmitem. Os autores trabalharam com uma ontologia de característica complexa de "Mina", que por definição, é uma escavação da terra para extração de minerais (*WordNet*).

Morais e Ambrósio (2007) explicam que a crescente complexidade dos objetos armazenados na Internet e o grande volume de dados exigem processos de recuperação cada vez mais sofisticados, apresentando novos desafios. De fato, com a explosão do número de documentos e de usuários na *WEB*, modelos para recuperação precisa de informações passaram a ser de grande importância. Em relação às ontologias, dois pontos principais devem ser considerados: Em primeiro lugar, é necessário trabalhar na criação de ferramentas que facilitem:

- Desenvolvimento de uma ontologia durante todo o ciclo de vida, incluindo: integração, fusão, reengenharia, avaliação de conteúdo, tradução para diferentes linguagens e formatos, troca de conteúdo com outras ferramentas,
- Gerenciamento de ontologias: gerenciamento de configuração e evolução de ontologias isoladas assim como de bibliotecas.
- Suporte de Ontologias: cronograma, documentação, técnicas avançadas para visualização do conteúdo da ontologia.
- Suporte metodológico para construção de ontologias.

Em segundo lugar, é necessário trabalhar na criação dos seguintes serviços:

- *Softwares* que ajudem a localizar a ontologia mais apropriada para uma dada aplicação;
- Métricas formais que comparam similaridades e distâncias semânticas entre termos da mesma ontologia ou em ontologias diferentes;

- *Softwares* que permitam incrementar, consistentes e seletivas atualizações da ontologia, a qual está sendo usada por uma dada aplicação;
- Módulos de *queries* para consultar ontologias;
- Acesso remoto para sistemas de bibliotecas de ontologias;
- *Softwares* que facilitem a integração da ontologia com sistemas de herança e bases de dados.

Noy e Mcguinness (2001) explicam que possibilitar a reutilização do conhecimento do domínio foi uma das forças motrizes por trás do recente aumento de pesquisa de ontologia. Por exemplo, modelos para muitos domínios diferentes precisam representar a noção de tempo. Esta representação inclui as noções de intervalos de tempo, pontos no tempo, medidas relativas de tempo, e assim por diante. Se um grupo de pesquisadores desenvolve tal ontologia em detalhes, outros podem simplesmente reutilizá-lo para seus domínios. Além disso, especificações explícitas de conhecimento de domínio são úteis para novos usuários que precisam aprender quais termos fazem parte de determinados domínios.

Quando uma ontologia de domínio é criada para contribuir com um SIG, ou seja, para tornar a informação geográfica mais útil, ela deve ser projetada baseada em atividades humanas diárias e nos seus respectivos espaços geográficos (KUHN, 2001).

2.6 RESUMO DA REVISÃO DE LITERATURA

A linguagem natural é estudada há décadas porque dentre outras aplicações desempenha um papel importante nos seres humanos, que ainda crianças desenvolvem uma percepção do espaço em que estão inseridos e começam a interagir entre si. De geração em geração a linguagem natural vai se espalhando através dos seres humanos, o que prova que se desenvolve num ambiente social e cultural dentro de comunidades. Dentro do contexto da linguagem natural falada e/ou escrita, percebe-se como o espaço pode ser representado de maneiras diferentes através de descrições de localização.

As descrições de localização contêm as mais diversas formas de termos de relações espaciais e não somente termos de relações espaciais topológicos. A

capacidade cognitiva espacial dos falantes e seu conhecimento do espaço variam de pessoa para pessoa, o que caracteriza a maneira como fazem descrições representando o espaço em que vivem e também como se orientam e se deslocam em suas rotinas diárias.

Os dois pontos principais nos estudos da cognição humana aplicada às linguagens naturais e mais especificamente na linguagem espacial, podem ser resumidas a duas hipóteses: “A linguagem espacial pode ser determinada pela categorização linguística do espaço?” Ou “O léxico da linguagem espacial restringe a forma e as categorias do pensamento?”. Estes pontos são responsáveis pela ligação da cognição humana com a linguagem natural e espacial, independente de qual seja a hipótese verdadeira.

A língua inglesa é uma das línguas naturais mais estudadas e exploradas, e em termos de linguagem espacial, existem estudos interlinguísticos que comparam outras línguas a ela no que diz respeito a alguns termos utilizados como relação espacial ou alguns verbos de movimento. Existem algumas preposições que foram estudadas em inglês e em espanhol, poucas preposições juntamente com o português europeu, mas com a língua portuguesa falada/escrita no Brasil estes estudos são escassos.

As relações espaciais fazem parte das descrições/diálogos dos seres humanos desde os primórdios quando havia a necessidade de se locomoverem e caçarem quando nômades. As relações espaciais foram e ainda são estudadas na cognição e linguística, nas ciências da computação e da informação geográfica. Dentre as classificações das relações espaciais, a mais explorada é a topológica, porém com os avanços tecnológicos e os estudos de PLN, existe a necessidade de estudos de mais termos de relações espaciais. A extração das relações espaciais ocorre a partir das expressões locativas, onde se descobrem quais são os *relata* e o *locata*. E nas descrições de localização realizadas em linguagem natural, cada pessoa tem a liberdade de utilizar quais termos ache necessário de acordo com sua percepção e sua cognição.

A semântica voltada para a cognição humana no que diz respeito ao contexto espacial vem ganhando espaço nos estudos atuais. Isto porque existe uma dificuldade de representação das descrições espaciais utilizadas no cotidiano das pessoas para o meio computacional. Como a semântica estabelece uma relação entre a linguagem e o pensamento e sugere um significado (triângulo semântico)

seu estudo aplicado às relações espaciais torna-se um desafio. A semântica vem contribuindo e facilitando a captura das informações na *WEB* em diversas áreas de aplicação. Os serviços *WEB* estão em constante mudança e melhoria com a aplicação da semântica e das ontologias.

Na última década estudos em similaridades vem contribuindo com a semântica, principalmente no que diz respeito à variedade de termos utilizados para a mesma situação. No contexto espacial, um termo pode ser usado para representar mais de uma situação e termos diferentes podem ser usados para uma mesma situação. Ou seja, as semelhanças e/ou diferenças encontradas através da similaridade contribuem diretamente na criação de ontologias.

3 MATERIAIS E MÉTODOS

3.1 MATERIAIS

Para a realização desta pesquisa foram utilizados um computador Intel Core i5 do laboratório de Cartografia do Programa de Pós-graduação em Ciências Geodésicas (PPGCG) da UFPR e um celular de propriedade da autora para gravação dos áudios de um dos testes. Também foi utilizado o pacote MS Office 2010 nas etapas de organização, estudo e análises dos dados coletados. Foram coletadas descrições escritas e/ou faladas em linguagem natural do português brasileiro, através de testes realizados *online* e em campo.

3.2 METODOLOGIA

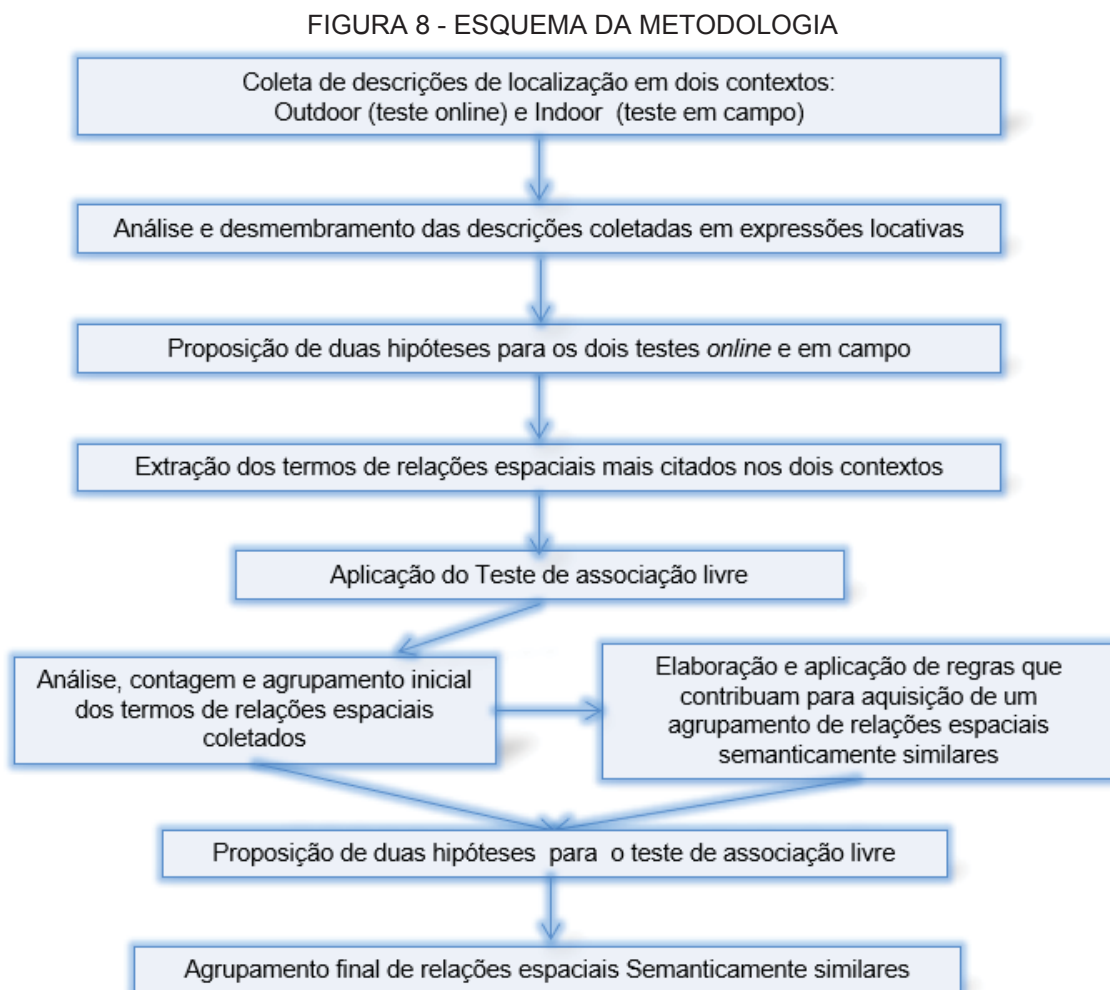
A organização da metodologia para atingir os objetivos específicos e chegar ao objetivo geral da pesquisa, de criar grupos semanticamente similares de termos de relações espaciais utilizados em linguagem natural que possam contribuir e ser utilizados em aplicações computacionais foi realizada conforme as etapas apresentadas na FIGURA 8.

A metodologia aplicada foi qualitativa e consistiu em testes com voluntários, tanto a partir de descrições de localização faladas como descrições escritas. O método qualitativo foi escolhido em função da pesquisa gerar e discutir entendimentos acerca de um tema de estudo específico e ter sido realizada com a coleta de dados a partir de experimentos com voluntários nativos da língua falada no Brasil em situações do dia a dia (SUCHAN; BREWER, 2000). O primeiro teste realizado consistiu na coleta de descrições de localização no contexto *outdoor* e foi aplicado *online*; o segundo teste foi no contexto *indoor* e foi aplicado *in loco*. Ambos os testes tiveram como foco a linguagem natural do português brasileiro para responder à pergunta: “Onde você se encontra localizado?”.

De posse das descrições realizadas nestes dois testes, foi possível realizar a etapa de análise e desmembramento das descrições em expressões locativas para extrair os termos de relações espaciais mais citados. Estes termos foram comparados entre os dois contextos e em seguida foram unidos em uma única tabela para serem extraídos os termos mais utilizados. Estes termos foram utilizados

no terceiro teste que consistiu na formulação e aplicação da associação livre de relações espaciais, sendo realizado *online*.

Com o resultado do terceiro teste foi possível analisar, contar e agrupar inicialmente os termos de relações espaciais coletados. De maneira paralela ocorreu a elaboração de cinco regras que contribuíram com a criação de um agrupamento de relações espaciais semanticamente similares.



FONTE: A autora (2021).

Os testes foram anônimos e apresentaram os termos de consentimento e concordância com os termos descritos pela UFPR para aplicação de testes (Apêndice 1). Também foram coletados alguns dados que caracterizaram os participantes, como: grau de instrução, área de formação, idade, frequência com que utilizam mapas, língua nativa. A proposição dos testes se baseou nos estudos das relações espaciais de Delazari et al. (2017) e Ferreira (2019) que trabalharam com

relações espaciais, Stock (2015) que trabalhou com identificação de padrões para a língua inglesa e Ragni, Tseden e Knauff (2007) que trabalharam com relações topológicas, e na consideração da localização baseada na percepção e na cognição humana de tudo que está próximo e/ou distante. Além de considerar que as pessoas passam por diversos locais diariamente e precisam se localizar com muita frequência, necessitando ter consciência sobre o local e as relações físicas existentes entre si e o meio ambiente (modo egocêntrico), juntamente com o contexto envolvido na situação (SCHMID et al., 2010).

3.2.1 Coleta de descrições de localização - *online*

A coleta de descrições de localização foi realizada por meio de um formulário *online*, com fotografias selecionadas a partir do *Google Street View*, de quatro regiões no território brasileiro e cinco regiões na Nova Zelândia. No território brasileiro foram selecionadas duas regiões urbanas centrais e duas regiões urbanas residenciais nas cidades de Curitiba no estado do Paraná e Uberlândia no estado de Minas Gerais. Na Nova Zelândia, as regiões selecionadas foram duas urbanas centrais, duas urbanas residenciais e uma não urbanizada, no município de Auckland. O contexto das fotografias foi o ambiente *outdoor*, ou seja, localização em ambientes externos e abertos como ruas e praças.

Pela diversidade na escolha das localizações das nove regiões, a realização deste teste foi *online* através do *Google Forms* e de forma randômica entre os participantes de modo que cada participante preenchesse o termo de consentimento e respondesse às perguntas somente uma vez. As localidades foram nomeadas por:

- BRA1: Praça Santos Dumond, bairro: Centro - Curitiba-PR,
- BRA2: Rua Eugênio Flôr, 1091, bairro: Abranches - Curitiba-PR,
- BRA3: Praça Tubal Vilela, bairro: Centro - Uberlândia-MG,
- BRA4: Felisberto Carrijo, 1059, bairro: Centro – Uberlândia-MG,
- NZ1: 309, Hibiscus Coast Hwy, Orewa – Auckland-NZ,
- NZ2: State Highway 1, Puhoi – Auckland-NZ,
- NZ3: 37, Wyndham St, Downtown – Auckland-NZ,
- NZ4: 79, Wainui Ave – Auckland-NZ
- NZ5: 147, Newton RD, Eden Terrace – Auckland-NZ

Cada uma das nove regiões foi composta por quatro fotografias que contemplaram as quatro direções a partir do ponto central, de modo a se ter todo o entorno dos locais representados em vista egocêntrica e para permitir que os participantes pudessem fazer suas descrições em linguagem natural escrita. As fotografias não continham os nomes de ruas, praças e identificação do local, fazendo com que os participantes realizassem as descrições de localização utilizando os pontos de referência que achassem convenientes. A situação criada e apresentada às pessoas trouxe o seguinte enunciado: “Imagine que você testemunhou um acidente de trânsito, precisa chamar o serviço de emergência e deve fornecer a localização do acidente de forma detalhada, sem citar nomes de ruas, somente com os pontos de referência, para ajudar o serviço a chegar ao local”. Esta situação baseia-se na pergunta “Onde você se encontra localizado?” e faz parte do projeto *Where am I?* que pretende transcrever descrições de localização em posições geográficas. Abaixo, as FIGURAS 9 e 10 ilustram dois grupos de fotografias capturadas de duas regiões do território brasileiro, sendo BRA1 numa região urbana central de Curitiba e BRA4 numa região urbana residencial de Uberlândia. E as FIGURAS 11 e 12 mostram dois grupos de fotografias capturadas de duas regiões da Nova Zelândia, sendo NZ3 numa região urbana central e NZ1 numa região não urbanizada de Auckland. Os demais grupos de fotografias encontram-se disponíveis no Apêndice 2.

FIGURA 9 - FOTOGRAFIAS DA ÁREA CENTRAL DE CURITIBA (BRA1)



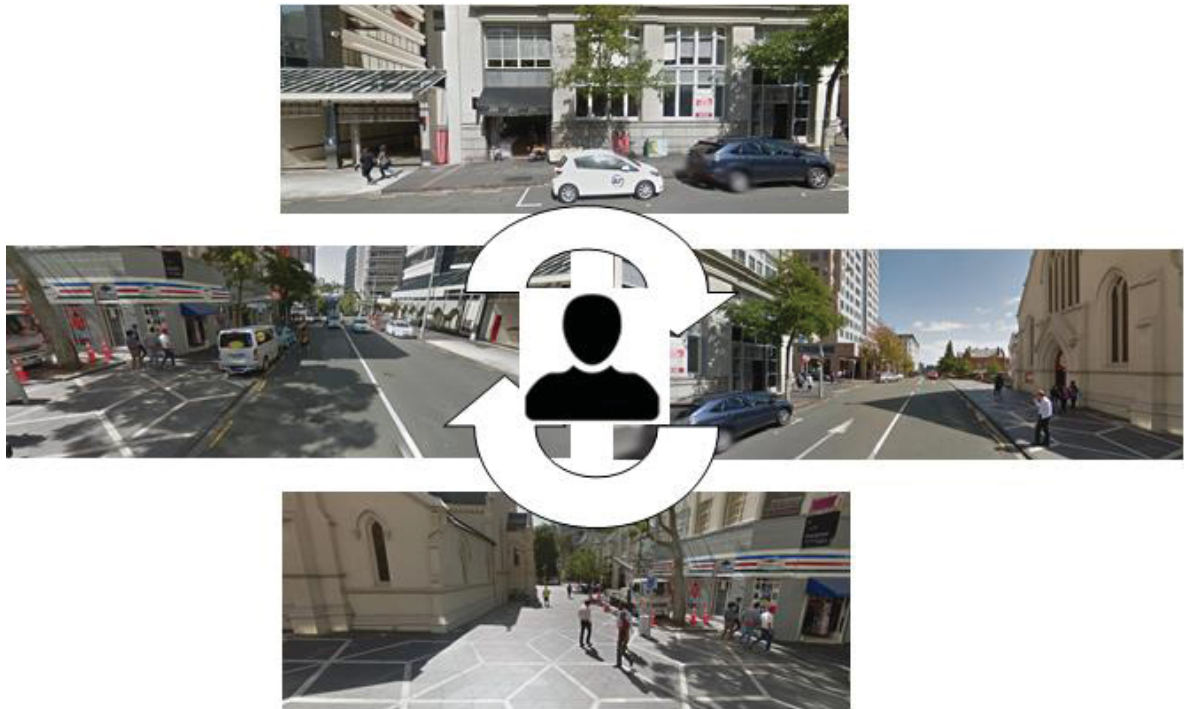
FONTE: A autora (2021).

FIGURA 10 - FOTOGRAFIAS DA ÁREA RESIDENCIAL DE UBERLÂNDIA (BRA4)



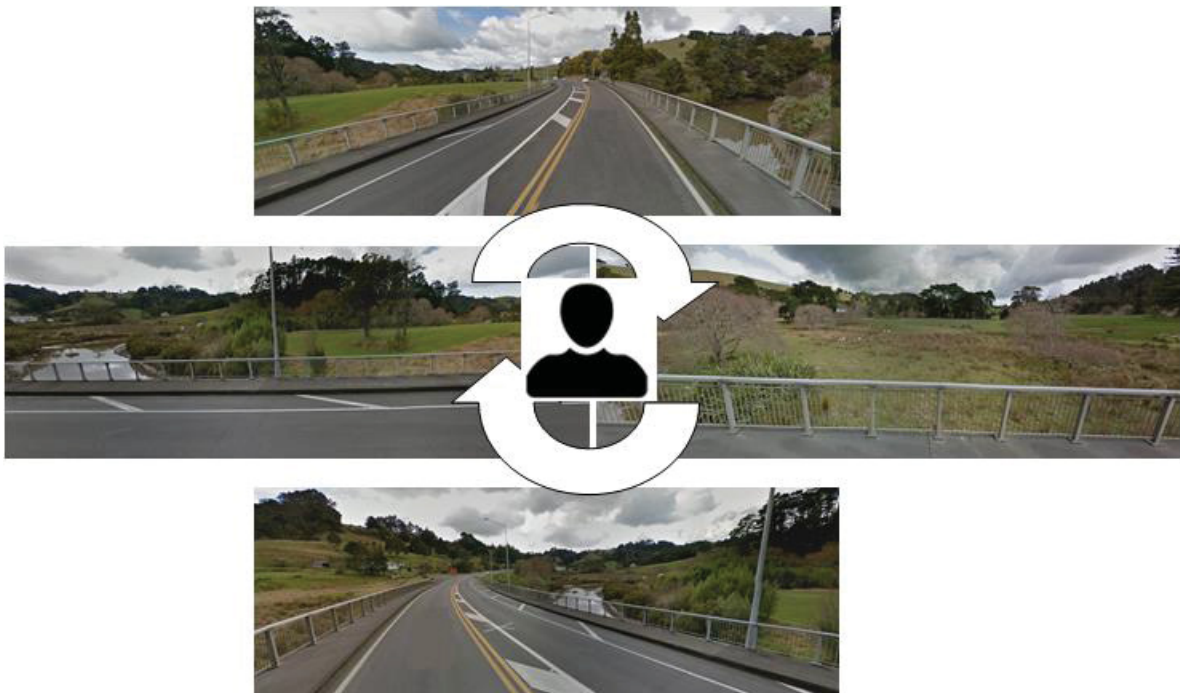
FONTE: A autora (2021).

FIGURA 11 - FOTOGRAFIAS DA ÁREA CENTRAL DA NOVA ZELÂNDIA (NZ3)



FONTE: A autora (2021).

FIGURA 12 - FOTOGRAFIAS DA ÁREA NÃO URBANIZADA DA NOVA ZELÂNDIA (NZ1)



FONTE: A autora (2021).

Após a coleta das descrições das localizações, iniciou-se o processamento dos dados que consistiu na análise e no desmembramento destas descrições em expressões locativas que são as sentenças que apresentam os *relata*, *locata* e os termos que representam as relações espaciais. Estas expressões locativas permitiram não só a extração dos termos de relações espaciais mais citados, como também os *locata* e os *relata* utilizados nas descrições para serem utilizados na etapa de análise das similaridades semânticas dos termos de relações espaciais.

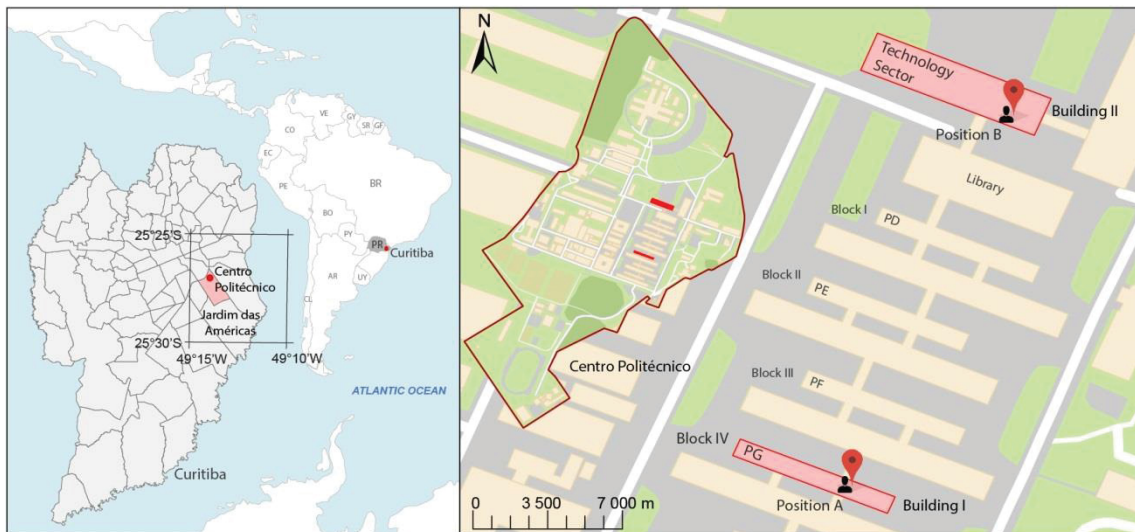
3.2.2 Coleta de descrições de localização - presencial

Esta etapa consistiu em um experimento com um questionário *online* e duas tarefas presenciais em ambientes *indoor*, com base na localização espacial dos participantes, com objetivo de coletar descrições de localização. Nestas tarefas foram coletadas as descrições verbais em língua portuguesa brasileira e que proporcionaram as análises de quais foram os termos de relações espaciais mais recorrentes para ambientes *indoor*.

As áreas para este teste foram predeterminadas em dois edifícios localizados no Campus III da Universidade Federal do Paraná no Centro Politécnico. O experimento foi realizado em duas etapas, sendo aplicado individualmente em cada área de estudo.

Os dois edifícios (FIGURA 13) que compuseram a área de estudo apresentam múltiplos andares e suas divisões estruturais são consideradas complexas por indivíduos que não conhecem o Campus. Além disso, dispõem de uma grande variedade informacional, como salas de aula, laboratórios didáticos, áreas de convívio comum, estabelecimentos comerciais, salas de departamentos/setores, salas de funcionários, auditórios, biblioteca e banheiros, que pode ser utilizada nas descrições de localização.

FIGURA 13 - LOCALIZAÇÃO DAS ÁREAS INDOOR



FONTE: A autora (2021).

O Edifício I foi projetado com o objetivo de desenvolver tarefas voltadas ao ensino, portanto seu contexto de uso é educacional e o Edifício II desenvolve tarefas administrativas (contexto administrativo). Nos experimentos, foi determinado um local em cada edifício (Posição A e B) para posicionar os participantes no ambiente de forma a se obter descrições de localização para serem extraídos os termos de relações espaciais.

Os testes foram divididos em duas etapas, sendo aplicados individualmente aos participantes em cada área de estudo (Edifícios I e II). Na Etapa 1, os participantes preencheram o termo de consentimento do teste e responderam um questionário *online* que se refere à sua identificação e caracterização (nome, idade, gênero, nível de formação). Em seguida, na mesma etapa discorreram sobre a frequência de visitas nos edifícios, área que o participante percorre nos edifícios e o nível de confiança em fornecer informações relacionadas aos edifícios para terceiros. Os participantes também autorizaram a gravação dos áudios.

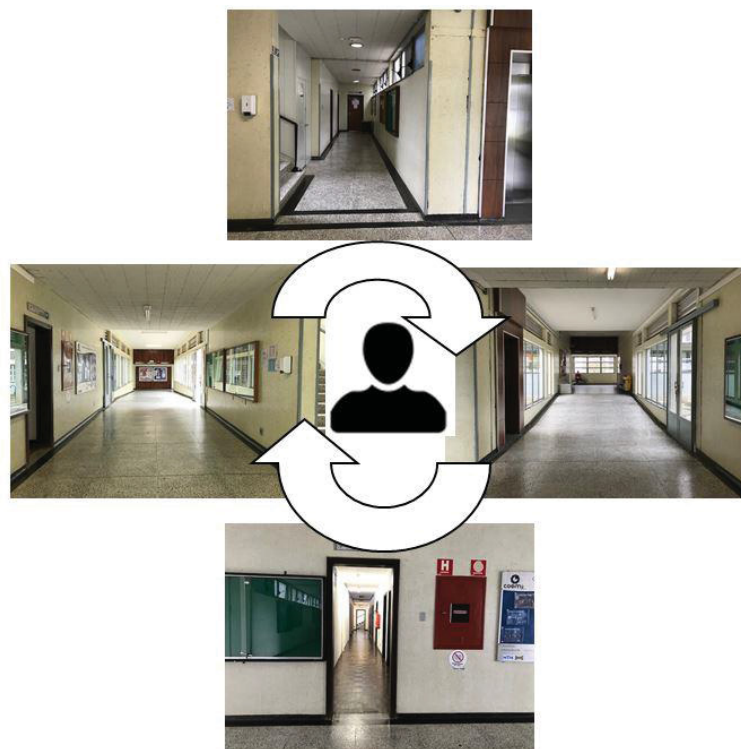
Na Etapa 2 os participantes foram levados aos pontos predeterminados nas duas áreas de estudo na Posição A e B (FIGURA 13). A posição A fica no Edifício I, no corredor do andar térreo do bloco de Engenharia Mecânica (FIGURA 14). E a posição B fica no Edifício II, no corredor do Salão Nobre no segundo andar no bloco da Administração (FIGURA 15).

Os participantes foram posicionados no centro dos dois corredores das áreas predeterminadas e apresentados à seguinte situação: Imagine que seu amigo

está lhe telefonando agora falando que acabou de chegar no Campus III e precisa encontrá-lo pessoalmente e lhe pergunta: “Onde você está localizado dentro do campus?”. Esta pergunta fez com que os participantes observassem os ambientes *indoor* onde estavam localizados e respondessem com duas descrições de localização. Estas descrições deveriam citar os elementos que os participantes consideravam importantes e as relações espaciais entre os próprios participantes e os elementos (modo egocêntrico) e/ou muitas vezes entre os próprios elementos existentes no ambiente (modo referenciado no objeto).

As respostas das duas perguntas foram gravadas *in loco* para que desta maneira não fosse perdida a fidelidade da língua natural falada. A gravação também fez com que a descrição fosse fluida e natural como acontece durante o processo de uma descrição em linguagem natural.

FIGURA 14 - POSIÇÃO A (EDIFÍCIO I) ANDAR TÉRREO DO BLOCO DE ENGENHARIA MECÂNICA



FONTE: A autora (2021).

FIGURA 15 - POSIÇÃO B (EDIFÍCIO II) SEGUNDO ANDAR DO BLOCO DA ADMINISTRAÇÃO



FONTE: A autora (2021).

As descrições de localização foram coletadas em áudio e gravadas num aparelho de celular. Os áudios foram armazenados e nomeados para serem vinculados aos questionários da Etapa 1. As descrições foram transcritas para uma planilha do Excel, onde foram analisadas e desmembradas em expressões locativas.

3.2.3 Análise e desmembramento das descrições de localização e extração dos termos de relações espaciais

As descrições de localização coletadas a partir dos dois testes foram transcritas para duas planilhas distintas do Excel, sendo uma para o contexto *outdoor* e outra para o *indoor*. As planilhas continham as descrições completas em uma coluna e o desmembramento das respectivas expressões locativas em outra coluna. Desta maneira, com o desmembramento em expressões locativas, foi possível realizar: 1) a extração e a contagem, 2) o agrupamento e a classificação dos termos utilizados como relações espaciais, 3) a verificação da existência de

novos termos utilizados como relações espaciais, e 4) a separação dos termos de relações espaciais a partir dos *locata* e dos *relata*.

3.2.4 Proposição de Hipóteses para os testes *online* e em campo

A partir da realização dos itens 1 ao 4 acima citados no item 3.2.3 foram levantadas as seguintes hipóteses para os dois testes:

1ª hipótese para os testes *online* e em campo: Se a coleta das descrições de localização em linguagem natural se deu em dois contextos diferentes – *outdoor* e *indoor* - e em situações geográficas com diferentes configurações, então, pelo grau de liberdade que a linguagem natural oferece, os termos de relações espaciais que se destacam são diferentes para cada tipo de ambiente.

2ª hipótese para os testes *online* e em campo: Se as pessoas fazem descrições de localização em linguagem natural em um determinado contexto e ambiente, de acordo com suas percepções do ambiente e do espaço e são livres para optar pelos elementos a serem relacionados, então elas utilizarão os termos de relações espaciais no referencial centrado nos elementos, ou seja, as relações espaciais se darão entre os elementos existentes no ambiente.

Além dos resultados das análises das duas hipóteses apresentadas, foi realizada a comparação dos dois testes executados, para geração de uma tabela com a união de todos os termos encontrados nos dois testes, sendo possível extrair os termos de relações espaciais mais mencionados nos dois contextos apresentados da língua portuguesa brasileira. Foram extraídos os termos que foram citados mais do que 10 vezes para serem utilizados no teste de associação livre de relações espaciais.

3.2.5 Teste de associação livre de relações espaciais - similaridade

O terceiro teste consistiu em coletar termos de relações espaciais, que as pessoas utilizam no seu cotidiano, a partir de sentenças de localização pré-elaboradas em linguagem natural. Após a coleta e através da análise e da contagem dos termos de relações espaciais foi possível realizar um agrupamento inicial destes termos. Também foram propostas e aplicadas regras que contribuíram com a criação e análise dos agrupamentos dos termos de relações espaciais e seus termos semanticamente similares que poderão ser utilizados computacionalmente na linguagem natural do português brasileiro.

O teste de associação livre de relações espaciais foi realizado *online* e na primeira parte, os participantes preencheram um questionário *online*, criado nos Formulários *Google*, contendo o termo de consentimento do experimento e perguntas sobre idioma nativo, nível de educação, curso, idade e sexo.

Na segunda parte, foram apresentadas sentenças onde o *locatum* foi designado como o sujeito “eu” para ter apenas a referência egocêntrica (KLATZKY, 1998) e os *relata* foram designados pensando-se em elementos que consistissem de uma base cartográfica, como por exemplo: rua, parque, rodoviária, prédios históricos.

A instrução deste teste de associação livre foi para que os participantes pensassem e escrevessem as palavras que tivessem o mesmo significado das relações espaciais destacadas nas sentenças, conforme quadro 1 abaixo. No enunciado do teste não foi escrito “similaridade” ou “sinônimo” porque isto induziria os participantes a ficarem focados em procurar os sinônimos e as similaridades das palavras em vez de pensarem diretamente nas relações espaciais que lhes viessem à cabeça.

Após a coleta dos dados, iniciou-se o processamento através da análise, contagem e agrupamento inicial dos termos de relações espaciais, de onde foram extraídos os termos citados pelos participantes.

QUADRO 1 - EXEMPLO DA ESTRUTURA DO TESTE DE ASSOCIAÇÃO DE PALAVRAS

Leia as frases abaixo e pense em palavras que tenham o mesmo significado das palavras que estão em letras maiúsculas.

Escreva as palavras que você pensou e que preenchem a lacuna do espaço em branco da frase

Faça isto para todas as frases.

Eu estou EM FRENTE À rodoviária.

Eu _____ rodoviária.

FONTE: AUTORA

3.2.6 Análise, contagem e agrupamento inicial dos termos de relações espaciais

Os dados foram exportados e organizados em uma planilha do Excel, onde as colunas continham os termos de relações espaciais das sentenças e as linhas continham todos os termos citados pelos participantes. Os termos citados foram agrupados inicialmente pelas semelhanças lexicais dos termos, mesmo quando estavam no singular ou plural, feminino ou masculino, ou quando havia apenas diferença nas preposições iniciais e/ou finais que acompanhavam o termo principal, e foram contabilizados. Se as respostas continham erros de escrita ou de gramática, mas entendia-se que faziam parte do mesmo agrupamento, também foram consideradas, por se tratar de linguagem natural.

3.2.7 Elaboração de regras para aquisição de termos semanticamente similares

A partir de um grupo de sentenças pré-definidas que estejam organizadas como expressões locativas, foram propostas e elaboradas regras para aquisição de termos de relações espaciais semanticamente similares. Portanto, de posse dos agrupamentos iniciais, conforme explicado acima no item 3.2.6 foram formuladas cinco regras para se obter os termos finais semanticamente similares:

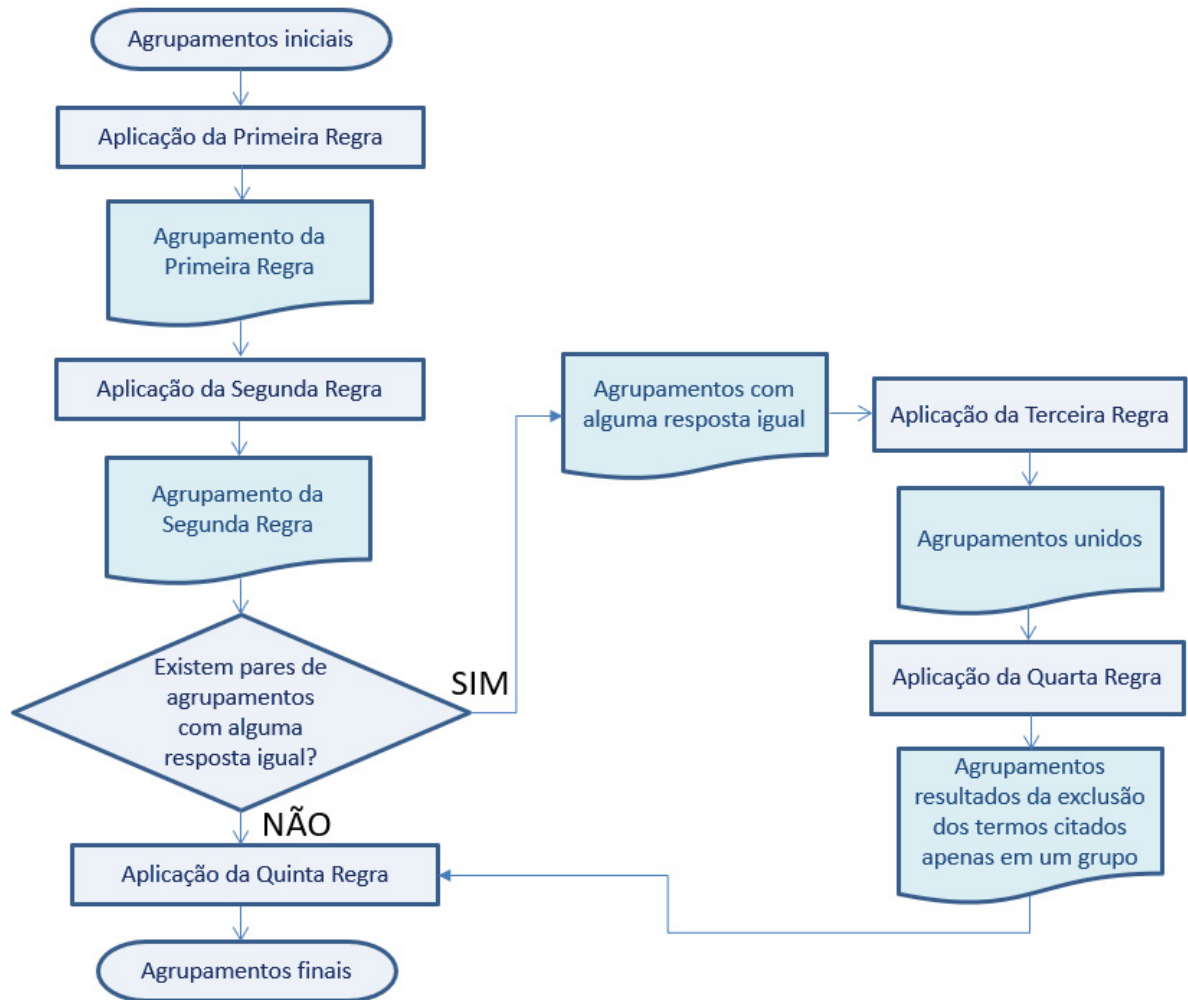
- ✓ Primeira regra: A partir dos agrupamentos iniciais, foram excluídos os termos citados menos de 1% do total de respostas de cada grupo de sentenças pré-definidas, para que estes termos não interferissem no

resultado de cada grupo. Pois são termos que não são comumente utilizados.

- ✓ Segunda regra: A partir do resultado da primeira regra, todos os termos de relações espaciais que compuseram as sentenças pré-definidas (neste caso os 14 termos principais) foram eliminados quando apareceram como respostas dos outros termos principais, ou seja, estes termos não deveriam aparecer como resposta em nenhum outro grupo. Esta regra colabora com a eliminação de termos de relações espaciais vagos que podem aparecer em vários grupos, atrapalhando o resultado das similaridades.
- ✓ Terceira regra: a partir do resultado da segunda regra, foi efetuada uma análise em pares de grupos para a verificação se houve algum grupo de termos com respostas iguais às respostas dos dois termos principais analisados. Se **sim**, estes dois grupos foram unidos. Ou seja, os grupos que tiveram respostas com termos de relações espaciais iguais caracterizam que podem ser termos similares; e passa-se para a quarta regra. Se **não** houve grupos de termos com respostas iguais nos pares de grupos analisados, passa-se para a quinta regra.
- ✓ Quarta regra: a partir da união dos grupos que obtiveram respostas com termos iguais (da terceira regra), foram considerados apenas os termos citados como resposta nos dois grupos. Os termos que foram citados em apenas um dos grupos unidos foram excluídos. Isto porque deveriam ficar somente os termos similares aos dois grupos que foram unidos.
- ✓ Quinta regra: dos agrupamentos recebidos da segunda regra ou da quarta regra, foram excluídos os termos que não apresentaram citações maiores do que 10% das citações do termo mais citado de cada grupo. Isto se justifica, por que podem ser termos que não sejam comumente utilizados. Como a finalidade é para aplicações computacionais, o ideal é que sejam utilizados os termos mais mencionados e que não gerem dúvidas quanto as suas similaridades semânticas e seu entendimento por todos os usuários.

Com a aplicação destas cinco regras, foi possível agrupar as relações espaciais e seus termos similares mais mencionados, conforme ilustra o fluxograma da FIGURA 16.

FIGURA 16 - FLUXOGRAMA DAS REGRAS



FONTE: A autora (2021).

3.2.8 Proposição de Hipóteses para o teste de associação livre

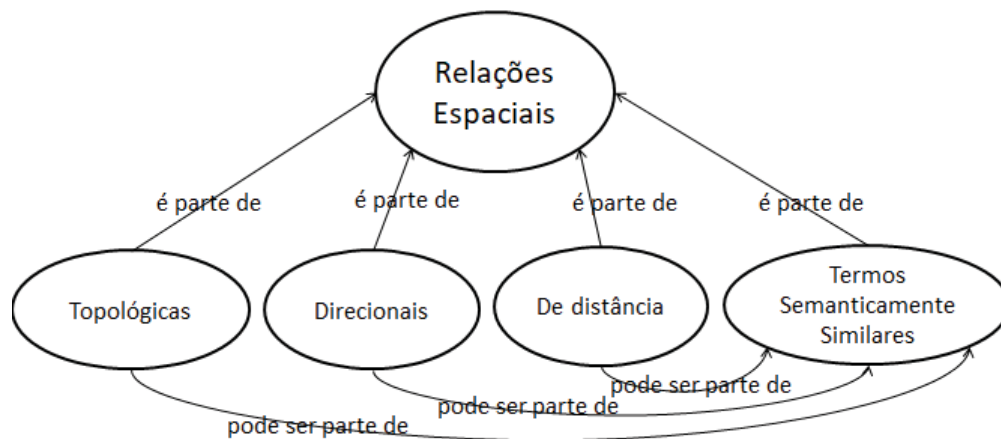
As hipóteses analisadas para este teste foram:

1ª hipótese para o teste de associação livre: Os termos semanticamente similares de relações espaciais encontrados através do mesmo *relatum*

e do mesmo *locatum* dos testes 3.2.1 e 3.2.2 coincidem com os termos encontrados no teste de associação livre.

2ª hipótese para o teste de associação livre: Se houver termos de relações espaciais que se comportem como semanticamente similares nas descrições, então pode ser realizado um agrupamento destes termos formando uma classe de termos semanticamente similares, conforme a FIGURA 17.

FIGURA 17 - AGRUPAMENTO DOS TERMOS DE RELAÇÕES ESPACIAIS



FONTE: A autora (2021).

3.2.9 Agrupamento final dos termos semanticamente similares

Os agrupamentos finais resultantes da aplicação das regras deverão analisados por um nativo da língua em questão. Os termos encontrados como similares poderão ser utilizados computacionalmente dependendo das aplicações a que se destinam. Por exemplo, em PNL numa aplicação de mineração de texto, podem ser utilizados todos os termos encontrados como similares; e em aplicações de geoprocessamento pode ser utilizado o termo considerado o mais similar.

4 RESULTADOS E ANÁLISES DOS RESULTADOS

Este capítulo apresenta os resultados e as análises obtidas nos três testes realizados para atingir o objetivo geral da pesquisa. Os participantes dos testes obrigatoriamente deveriam ser nativos da língua portuguesa brasileira, a fim de se obter os termos de relações espaciais utilizados em linguagem natural. Os resultados e as análises foram descritos de acordo com o esquema da metodologia.

4.1 COLETA DE DESCRIÇÕES DE LOCALIZAÇÃO - *ONLINE*

O teste *online* consistiu na definição de áreas em diferentes cidades com configurações urbanas e rurais distintas visando a coleta de descrições de localização em linguagem natural do português brasileiro no contexto *outdoor*. Foi realizado com voluntários, aplicado *online* e de forma randômica, de modo que cada voluntário respondesse às perguntas somente uma vez.

A realização deste primeiro teste foi a partir da seleção de quatro regiões no território brasileiro, sendo uma região urbana central e uma região urbana residencial nas cidades de Curitiba e Uberlândia, que ocorreu com a coleta dos dados no período de abril e maio de 2016. E a partir da seleção de cinco regiões na Nova Zelândia, sendo duas urbanas centrais, duas urbanas residenciais e uma rural, que ocorreu no período de março a maio de 2017.

A partir do total de descrições coletado foi realizado o desmembramento em expressões locativas para que fosse possível realizar: 1) a extração e a contagem dos termos utilizados como relações espaciais e seus respectivos *relata* e *locata*, 2) o agrupamento e a classificação dos termos utilizados como relações espaciais, 3) a verificação da existência de novos termos utilizados como relações espaciais, e 4) a separação dos termos de relações espaciais a partir dos mesmos *locata* e *relata*.

4.1.1 Coleta das descrições de localização e desmembramento em frases locativas

A coleta dos dados das áreas do Brasil ocorreu no período de abril e maio de 2016 e das áreas da Nova Zelândia no período de março a maio de 2017. Foram coletadas e organizadas numa planilha do Excel um total de 121 descrições de localização, sendo 64 descrições referentes às quatro áreas do território brasileiro e

57 referentes às cinco áreas do território neozelandês, todas realizadas por nativos da língua portuguesa brasileira. As descrições encontram-se na base de dados científicos (<http://dx.doi.org/10.5380/bdc/65>).

O primeiro passo foi a organização das descrições em pastas separadas por localidades, onde as 121 descrições foram desmembradas em 557 expressões locativas, ou seja, destas descrições foram utilizados 557 termos de relações espaciais (TABELA 1). O desmembramento das descrições ocorreu pela necessidade de se entender quais foram os termos de relações espaciais utilizados entre o observador (*locatum*) e os elementos (*relata*), ou entre um elemento referenciado (*relatum*) e o elemento localizado (*locatum*).

TABELA 1 - TOTAL DE TERMOS DE RELAÇÕES ESPACIAIS CONTEXTO *OUTDOOR*

Localidades	Termos de relações espaciais
BRA1	94
BRA2	57
BRA3	79
BRA4	91
NZ1	63
NZ2	58
NZ3	51
NZ4	21
NZ5	43
TOTAL	557

FONTE: A autora (2021).

Cada descrição teve que ser entendida e analisada primeiramente no aspecto geral e depois parte a parte para serem criadas as expressões locativas. As expressões locativas continham os termos que as pessoas utilizaram como relações espaciais, os seus respectivos *relata* e *locata*. Desta maneira, pôde-se ter uma ideia de como as pessoas utilizam os termos de relações espaciais no seu cotidiano em linguagem natural.

Para cada relação espacial, um *relatum* (feição geográfica/objeto) foi identificado. Se uma única relação espacial se referisse a múltiplas feições geográficas/objetos, por exemplo, “A casa está em frente à igreja e aos correios.”, foram criados tantos pares separados de relação espacial quantos necessários. No caso do exemplo, foram criadas duas expressões locativas para o *locatum* casa, tendo dois *relata* igreja e correios. Em alguns casos, a expressão envolveu a utilização de um pronome para alguma feição/objeto, por exemplo, “Estou localizado

em uma esquina de praça, ela fica em frente a um bar chamado TXAPELA e próxima a um hotel chamado LAOS.”, onde “ela” se refere à praça. Houve vários casos em que o *relatum* para uma relação espacial era ambíguo, por exemplo, “Ocorreu um acidente ao lado da estátua, ao lado do banco.”. Neste caso não é imediatamente claro o que está ao lado do banco: o acidente ou a estátua. A presença da vírgula sugere a primeira hipótese e como as descrições foram analisadas caso a caso juntamente com as fotografias, pôde-se confirmar o *relatum* adequado em casos como esse.

Em alguns casos, o *locatum* não foi mencionado, mas estava implícito na descrição, por exemplo, as expressões “perto do parque, ao lado da estátua” consistem em dois pares de relação espacial, ambos com *locatum* implícito que é o observador. É possível interpretar isso porque muitas vezes o verbo estava na primeira pessoa do singular, por exemplo, “estou perto do parque, estou ao lado da estátua”; além do que a gramática da língua portuguesa brasileira utiliza sentenças com sujeito oculto em muitas situações. Outro tipo de expressão comum que omite o *locatum* é quando se tem a utilização dos termos “há” e “tem”, por exemplo, “Há/tem estacionamento para motos.”, esta expressão indica que existem lugares de estacionamento para motos no local. Em casos como estes, foi considerado que existe a relação espacial e o estacionamento de motos como *relatum*, porém sem *locatum*.

4.1.2 Extração e contagem dos termos de relações espaciais e seus respectivos

relata e locata

Os resultados da extração e da contagem dos termos de relações espaciais utilizados e seus respectivos *relata* e *locata* foram apresentados em tabelas por localidades. Abaixo segue um exemplo relacionado à área urbana central de Curitiba (BRA1), de uma descrição de localização realizada por um voluntário e de como foi realizado o desmembramento de todas as descrições em expressões locativas. Todas as descrições referentes a todas as localidades seguiram a mesma metodologia.

“Estou localizado em uma esquina de praça, que fica em frente a um bar chamado TXAPELA e próxima a um hotel chamado LAOS.”

O desmembramento da descrição acima resultou em três expressões locativas e na extração de três termos utilizados como relações espaciais, três *locata* e três *relata*:

1. Estou localizado ***em uma*** esquina de praça. Nesta expressão locativa, o termo de relação espacial é “em uma”, o *locatum* foi implícito dado pelo sujeito “eu” e o *relatum* foi a “esquina de praça”, sendo que a “esquina de” é um termo que especifica o *relatum*.
2. Praça que fica ***em frente a*** um bar chamado TXAPELA. Nesta expressão locativa, o termo de relação espacial é “em frente a”, o *locatum* é a “praça” e o *relatum* é o “bar”.
3. Praça ***próxima a*** um hotel chamado LAOS. Nesta frase locativa, o termo de relação espacial é “próxima a”, o *locatum* para este caso é a “praça” e o *relatum* é o “hotel”.

A língua portuguesa brasileira aceita sujeitos ocultos para os verbos, portanto algumas vezes os voluntários utilizaram nas descrições o sujeito “eu” e/ou algum elemento de forma implícita. Estes sujeitos e estes elementos foram denominados nas análises como “termos implícitos” e foram contados de acordo com sua utilização.

Este modelo de desmembramento foi aplicado para as nove localidades, que resultaram em nove tabelas com as colunas *Locata*, *Relata* e Termos de relações espaciais, onde os termos implícitos foram anotados com a palavra “implícito”. E na coluna dos termos de relações espaciais, os termos (l) e (r) representam os *locata* e *relata* respectivamente, as nove tabelas encontram-se na base de dados científicos (<http://dx.doi.org/10.5380/bdc/73>).

Para a primeira localidade analisada BRA1, o desmembramento resultou em 94 termos de relações espaciais com os seus respectivos *locata* e *relata*. A TABELA 2 apresenta parte destes resultados.

TABELA 2 - DESMEMBRAMENTO RELAÇÕES ESPACIAIS - LOCALIDADE BRA1

Locata	Relata	Termos de relações espaciais
eu	cruzamento	(l) me encontro no (r)
rua menor de paralelepípedos	cruzamento	(r) com (l)
edifício de características coloniais	praça	(l) em frente à (r)
edifício de características coloniais	rua asfaltada	(l) do outro lado da (r)
acidente	via asfaltada e via secundária de terra	(l) aconteceu entre (r) e (r2)
acidente	esquina do supermercado Pop	(l) na (r)
mercado	construção amarela de dois andares	(l) fica na parte inferior (r)
pinheiro	mercado	ao lado direito do (r) há (l)
eu	esquina de praça	(l) estou localizado em uma (r)
esquina de praça	bar chamado TXPELA	(l) fica em frente a (r)
esquina de praça	hotel chamado LAOS	(l) próxima a (r)
acidente	esquina	(l) numa (r)
acidente	local	(l) em um (r)
prédios históricos	local	(r) com (l)
local	praça	(l) próximo á (r)
rua de paralelepípedo	sentido leste-oeste	no (r) há (l)
duas edificações de cor amarela	prédios	dentre (r) estão (l)
comércio	duas edificações de cor amarela	entre (r plural) um (l)
local	central	(l) deve ser (r)
implícito - acidente	rua sentido norte-sul	(l implícito-acidente) na (r)
rua sentido norte-sul	estabelecimento comercial de cor amarela	(l) uns metros acima do (r)

FONTE: A autora (2021).

Podem ser visualizados espacialmente, na forma de mapa, os *locata* e *relata* que foram mencionados nas descrições de localização. Quando se observa a FIGURA 18 percebe-se que os *locata* mais utilizados para a localidade BRA1, representados por elementos pontual foram: “eu” e “acidente” que estão representados no mesmo lugar, ou seja, no cruzamento, seguidos dos *locata* “estátua e árvores”; por elementos lineares foram mencionadas as duas ruas do local, porém com a diferença de que as pessoas preferiram mencionar como elemento localizado a rua estreita e de paralelepípedos; e em áreas, aparecem a praça e o comércio na esquina como os mais citados.

FIGURA 18 – VISUALIZAÇÃO ESPACIAL DOS LOCATA – BRA1



FONTE: A autora (2021).

A espacialização dos relata da localidade BRA1 na FIGURA 19 mostra que o “cruzamento” foi a feição pontual mais utilizada, seguida da rua asfaltada e de sentido único e da praça. Isto prova que as pessoas observam o local e escolhem menos feições como *relatum* e ficam mais livres para mencionar outros elementos como *locata*, visto a maior preferência por três feições como relata.

De acordo com Schmidt (2012), Vinson (1999) e Werner et al. (1997) a praça, e a rua asfaltada de sentido único se sobressaem no local se destacando dos demais elementos do entorno, e podem ser classificados como pontos de referência visuo-espaciais porque suas características são contrastantes no local (LYNCH, 1960; SORROWS; HIRTLE, 1999).

Os demais elementos mencionados poucas vezes (bancos, banca de jornais, ponto de ônibus, bandeiras, prédios altos, hotel) podem ser considerados pontos de interesse e serviram para alguns participantes, mas foram irrelevantes para outros (FANG; LI; SHAW, 2015; RICHTER; WINTER, 2014).

FIGURA 19 – VISUALIZAÇÃO ESPACIAL DOS RELATA – BRA1



FONTE: A autora (2021).

4.1.3 Agrupamento e a classificação dos termos de relações espaciais

O agrupamento dos termos de relações espaciais foi realizado em seis etapas:

1. Gênero: verificou-se a existência de termos que remetiam ao gênero feminino e masculino. Como por exemplo, nos casos “(l) em frente à (r)” e “(l) em frente ao (r)” onde os termos de relações espaciais foram utilizados remetendo-se a um *relatum* do gênero feminino, dado por “à” ou pelo gênero masculino representado por “ao”.
2. Singular e plural: verificaram-se os termos citados no singular e no plural, como “(l) em direção à (r)” onde o “à” é singular e “(l) em direção aos (r)” onde “aos” é plural.
3. Preposições iniciais e finais: junção dos termos que apresentavam diferentes preposições no início e no final dos termos de relação espacial, como por exemplo, “(l) do lado de (r)” onde “do” é a preposição “de” combinado com o artigo definido “o” no início do termo; e “de” é a

preposição do final do termo; e “(l) ao lado da (r)” onde “ao” é a preposição “a” combinado com o artigo definido “o” no início e “da” é a preposição “de” combinado com o artigo definido “a”.

4. Verbos com termos de relação espacial: agrupamento pelos verbos que acompanharam os termos de relações espaciais, pois muitos termos de relações espaciais foram utilizados nas descrições em linguagem natural acompanhados de verbos, por exemplo: “(l implícito - eu) estou situado em frente a (r)” ou “(l) ocorreu em frente a (r)”.
5. Relações espaciais: agrupamento por vocábulos dos termos de relações espaciais, ou seja, pelos próprios termos principais, como por exemplo, em frente, próximo, perto, atrás, ao lado.
6. Expressões com verbos e sem relações espaciais: possui, consigo ver, há.

Os agrupamentos dos termos de relações espaciais foram realizados a partir de todas as expressões locativas das nove localidades e seguindo as seis etapas acima. A TABELA 3 mostra como exemplo, o agrupamento resultante para o termo de relação espacial “em frente”. Este agrupamento apresentou como resultado 33 maneiras de utilização deste termo. Todos os demais termos foram agrupados seguindo o mesmo raciocínio e encontram-se na base de dados científicos (<http://dx.doi.org/10.5380/bdc/66>).

A classificação dos termos de relações espaciais foi realizada a partir das classes existentes em estudos sobre relações espaciais: topológica, direcional, projetiva ou cardinal; ou quando descreviam distância, adjacência ou orientação (BRUNS; EGENHOFER, 1996; COVENTRY; GARROD, 2004; KEMMERER, 2006; TENBRINK, 2017)

TABELA 3 - AGRUPAMENTO DO TERMO DE RELAÇÃO ESPACIAL “EM FRENTE” DO TESTE ONLINE

Termo de relação espacial	Agrupamento do termo de relação espacial “em frente” - <i>OUTDOOR</i>
em frente	(l implícito - acidente) em frente a (r), (l implícito - acidente) em frente á (r), (l implícito - acidente) em frente ao (r), (l implícito - acidente) na frente de (r)
	(l) em frente a (r), (l) em frente à (r), (l) em frente ao (r), (l) na frente de (r), (l) em frente da (r), (l) na frente da (r)
	(l implícito - eu) estou em frente a (r), (l implícito - eu) estou em frente à (r), (l implícito - eu) estou em frente as (r), (l implícito - eu) estou em frente ao (r), (l implícito - eu) estou na frente do (r)
	(l) estou em frente (r), (l) estou em frente à (r), (l) estou em frente ao (r)
	(l) fica em frente a (r), (l) fica em frente ao (r)
	(l) ocorreu em frente a (r)
	(l) que fica em frente a (r)
	em frente (r) há (l), em frente a (r) há (l)
	(l implícito - acidente) bem em frente a (r)
	(l implícito - eu) me encontro em frente à (r), (l implícito - eu) me encontro em frente ao (r)
	(l) aconteceu em frente ao (r)
	(l) estão em frente a (r), (l) está em frente ao (r)
	(l implícito - eu) encontro-me em frente à (r)
	(l implícito - eu) estou situado em frente a (r)
	(l) em frente (r implícito - prédio histórico)
	(l) estiver parada na frente da (r)
	(l) estou localizado em frente a (r)
	(l) frente (r)
	(r) tem (l) na frente
	em frente a essa (r) existem (l)
	em frente a esta (r) tem (l)
	existe (l) em frente a (r)
	há (l) em frente (r implícito - Escola Bueno Brandão)
	há ainda em frente essa (r) (l)
	quase em frente a (r) há (l)
	quase em frente do (r) temos (l)
	tem (l) na frente da (r)
	(l) em frente (r implícito - acidente)
	(l) que passa em frente da (r)
	tem (l) bem em frente ao (r)
	(l) está bem em frente a (r)
	(r) com (l) na frente
há (l) na frente da (r)	

FONTE: A autora.

O resultado do agrupamento e da classificação de todos os termos encontrados de relações espaciais no teste *online* de contexto *outdoor* está na TABELA 4. Esta tabela apresenta os 557 termos de relações espaciais agrupados por gênero, plural, preposição, verbos e distribuídos em 60 grupos de relações espaciais com suas respectivas classificações.

TABELA 4 - AGRUPAMENTO E CLASSIFICAÇÃO DOS TERMOS DE RELAÇÕES ESPACIAIS DO TESTE *ONLINE* (continua)

Classificação	Agrupamento de termos de relações espaciais	Total
Projetiva	em frente, na frente	96
	ao lado de, do lado da	22
	lado direito, lado esquerdo	13
	do outro lado da	13
	em um lado, de um lado	8
	direita, esquerda	6
	à frente	6
	lado par, lado ímpar	4
	de frente	3
	de frente	3
	em ambos os lados, em ambas as laterais	3
	após	3
	abaixo	3
	em cima	3
	sobre	3
	por cima	3
	adiante	2
	na outra	2
	sob	2
	ao final	1
	lado contrário	1
	oposta	1
	antes	1
	frente a frente	1
	para o lado	1
	um lado da	1
	ao/no fundo	2
	atrás da	1
	depois	1
	cujo fim	1
cujo fundo	1	
inferior	1	
acima de	1	
Topológica	no, na	56
	em um, em uma, em que	43
	numa, num	10
	entre	8
	neste, nesta	7
	lado de dentro, lado de fora	2
	em	2
dentre	1	
Distância	próximo	41
	perto do, perto de	14
	x distancia	4
	x metros acima do	2
	nas proximidades	2
	nas imediações	1

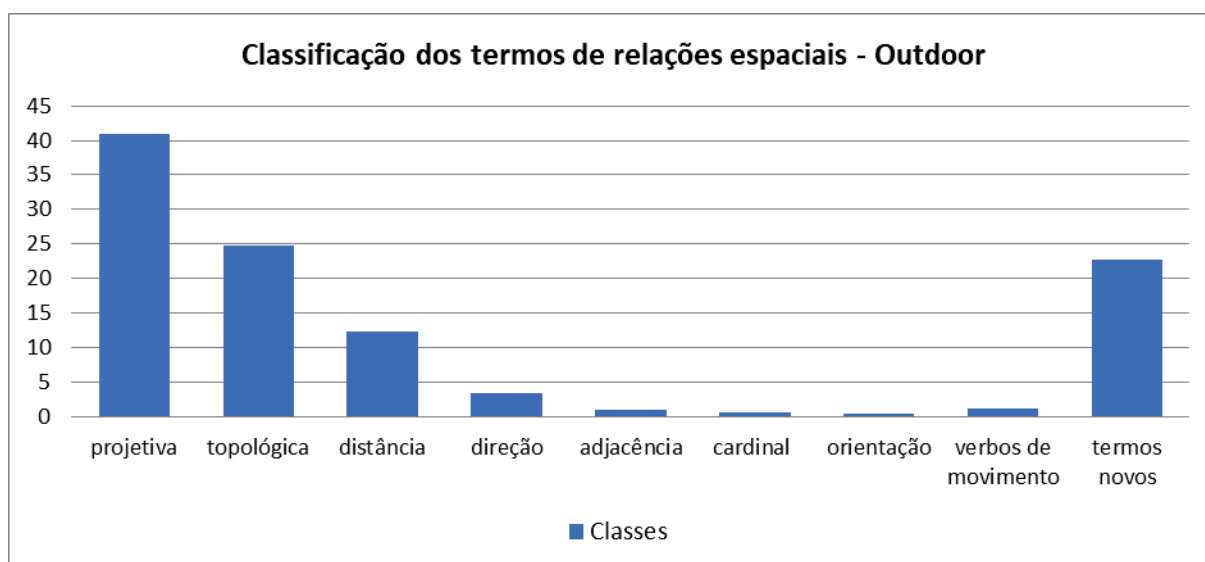
TABELA 4 - AGRUPAMENTO E CLASSIFICAÇÃO DOS TERMOS DE RELAÇÕES ESPACIAIS DO TESTE *ONLINE* (conclusão)

Classificação	Agrupamento de termos de relações espaciais	Total
Direção	em direção a, em direção ao	6
	ao longo	2
	vai do xx para o yy	1
	no sentido	1
	Acesso ao	1
	ao redor	2
	direção norte, direção sul	2
	em outra direção	1
Adjacência	em volta	1
	junto a	3
	adjacente	1
Cardinal	anexo	1
	ao sul, ao norte	3
Orientação	perpendicular	2
Verbos de movimento	passa, termina, chega, cruza, divide	6
Termos novos	com	79
	verbos	39
TOTAL		557

FONTE: A autora (2021).

Percebe-se pelo GRÁFICO 1 abaixo que para a língua portuguesa brasileira a classe projetiva ficou com 41% das citações se sobressaindo em relação às demais classes:

GRÁFICO 1 - CLASSIFICAÇÃO DOS TERMOS DE RELAÇÕES ESPACIAIS TESTE *ONLINE*



FONTE: A autora (2021).

A TABELA 5 diz respeito ao agrupamento dos 39 verbos citados como relações espaciais que estão agrupados como “verbos” na TABELA 4.

TABELA 5 - AGRUPAMENTO DOS VERBOS DO TESTE ONLINE

Agrupamento dos Verbos	Total
(r) tem (l), (r) que tem (l), tem (l), quando termina a lateral da (r), tem (l)	10
(r) possui (l), (r) possui (l) em ambos os sentidos	8
(l) consigo ver (r), (l) vejo (r), (l) verei (r), já pode se ver um (r) o qual tem um (l)	4
(r) há (l), há (l), (r) onde há (l), (r) somente há (l)	4
(r) que contém (l), (r) contém (l)	3
(l) deve ser (r), (l) que parece (r), no que parece ser uma (r)	3
(r) onde existe (l), (r) existem (l)	2
(l) olhar para (r)	1
(l) pode ser avistada desta (r)	1
(r) apresentam (l)	1
(r) onde estão (l1) e (l2)	1
(l) onde não se consegue visualizar o que há para dentro dos (r)	1
TOTAL	39

FONTE: A autora (2021).

4.1.4 Novos termos utilizados como relações espaciais

Os resultados mostram que os voluntários utilizaram novos termos como relações espaciais em suas descrições como, por exemplo, o termo “com” e alguns verbos.

O termo “com” foi utilizado nas descrições de localização quando estava caracterizando o *locatum* e muitas vezes ele pôde ser utilizado como relação espacial. Resumindo, o termo “com” foi mencionado pelos voluntários 79 vezes, porém foi utilizado 36 vezes somente para caracterizar elementos, ou seja, sem contribuição espacial. E expressões desta maneira ocorreram em todas as localidades, conforme exemplos abaixo:

- Prédio histórico grande e muito bem conservado com bandeiras hasteadas. (BRA1)
- Supermercado Super Pop com símbolo da Pepsi. (BRA2)
- Edifício com fachada trapezoidal com janelas e portas azuis. (BRA3)
- Prédio branco em reforma com elevador do lado de fora. (BRA4)
- Casa verde com um carro vermelho na garagem. (NZ1)
- Fazenda com alguns gados pastando. (NZ2)
- Igreja com arquitetura bizantina/gótica. (NZ3)
- Empresa chamada *Water Care* com o número 12. (NZ4)

- Edifício na cor verde com uma loja no térreo. (NZ5)

Para a finalidade de relações espaciais estas 36 utilizações não foram computadas, portanto, o termo “com” ficou com a contagem final de 43 utilizações como relação espacial. Ainda um número considerado alto de citações e ocupando o quarto lugar na classificação geral dos termos, concordando com Kleppa (2008) que comentou que esta é uma das preposições mais utilizadas e discordando quando o autor afirma que a preposição “com” não tem uma semântica relacionada ao espaço.

De acordo com a classificação que Egorova (2018) aplicou no seu estudo, a preposição “com” pode ser inserida na classe que descreve os pontos de referência visuais. Segundo a autora, esta classe mostra-se bastante utilizada nas descrições de rota em contextos urbanos conforme provaram os resultados de utilização da preposição “com” no contexto *outdoor* deste artigo.

Abaixo seguem alguns exemplos desta preposição que foram utilizados como relações espaciais.

- “Acidente em uma esquina, em um local com prédios históricos, próximo à uma praça” (BRA1). Onde a expressão locativa “local com prédios históricos” representa além da característica do local, porque se pode filtrar a localização como sendo um local onde existem prédios históricos.
- “Estou em frente ao supermercado Superpop, em um cruzamento com 3 saídas, interseção em forma de T” (BRA2). Onde a caracterização do cruzamento contribui com a localização por ser diferenciada dos demais cruzamentos.
- “Rua com duas faixas e sentido único, acidente próximo a um hotel de frente a uma praça onde existe uma rotatória com um ponto de táxi” (BRA3). A rotatória com a explicação de que nela existe um ponto de táxi a diferencia de outras rotatórias e colabora com a localização.
- “Área residencial, pista de mão única com duas faixas de rolamento, nas proximidades do número 1060, primeira à esquerda após o Sítio da Trindade vindo direção mercado de casa amarela” (BRA4). Como o

bairro é residencial, a característica de uma pista de mão única com duas faixas é um diferencial que contribui com a localização.

- “Me encontro na avenida a beira mar, em frente ao Starring, local de cor vermelha e amarela, com estacionamentos para motocicletas e uma calçada de *petit pavê*, um lado da rua é comercial e o outro residencial” (NZ1). As expressões “Starring-local com estacionamento para motocicletas” e “Starring-local com uma calçada de *petit pavê*” diferenciam o local na avenida beira mar.
- “Solicito um socorro de ambulância, não sei informar o nome da rua, mas me encontro em Frente ao comércio City Mart localizado entre em uma rua, que contém uma praça e uma possível igreja/catedral com uma grande calçada e praça de passagens de pedestres” (NZ3). A igreja/catedral com uma grande calçada e praça de passagem de pedestre é o diferencial da região, contribuindo com a localização.
- “Final de uma rua com declive acentuado” (NZ4).
- “Acidente em uma rua de mão dupla, via em T, com prédios antigos nas imediações” (NZ5).

O termo “com” foi classificado como topológico porquê de acordo com sua utilização como relação espacial, ele significa que o *locatum* está inserido ou pertence ao *relatum*, como nos exemplos acima, das descrições 10 a 17 e seu agrupamento resultou na TABELA 6.

TABELA 6 - RELAÇÕES ESPACIAIS UTILIZANDO O TERMO “COM”

Com	Total
(r) com (l)	40
(r) que se parece com (l)	1
(r) apenas com (l)	1
(r) com outra (l)	1

FONTE: A autora (2021).

Para os verbos, o agrupamento foi realizado pelo tipo do verbo que os voluntários utilizaram como relação espacial nas suas descrições. Todas as descrições que continham um determinado tipo de verbo, fosse ele mencionado sozinho ou em uma relação com mais palavras, foram agrupadas. Por exemplo, o

verbo “ter” além de ter sido utilizado sozinho, foi utilizado com os vocábulos: “que” e “quando termina a lateral da” (TABELA 5).

Da maneira como foram utilizados, eles ainda foram agrupados pelos seus sinônimos e obedecendo às regras da língua portuguesa brasileira, como por exemplo:

- Ter (mencionado 10 vezes), possuir (mencionado 8 vezes) e conter (mencionado 3 vezes),
- Existir (mencionado 2 vezes) e haver (mencionado 4 vezes).

Para que os verbos entrassem nas classificações existentes de relações espaciais (topológica, direcional, projetiva, cardinal, distância, adjacência ou orientação), eles foram analisados nas sentenças e foram trocados por relações espaciais que não mudassem os seus significados, ou seja, buscaram-se termos que fossem semanticamente similares para estes verbos quando possível. Como por exemplo:

- “consigo ver, vejo, verei, pode ser avistada, pode se ver” podem ser substituídos pelo termo “à frente” nas expressões locativas e podem ser colocados na classe projetiva. Seguem alguns exemplos:

QUADRO 2 - SUBSTITUIÇÕES PELO TERMO "À FRETE"

Expressões locativas	Substituição
Consigo ver um prédio em construção.	À (minha) frente tem um prédio em construção.
Vejo muitas motos.	À (minha) frente tem muitas motos.
Verei um declive.	À (minha) frente tem um declive.
A praia pode ser avistada desta rua.	A praia está à frente desta rua.
Pode se ver um sobrado grande e amarelo.	À frente tem um sobrado grande e amarelo.

FONTE: A autora (2021).

- “olhar para” pode ser substituído por “em direção” e foi classificado como de direção.

QUADRO 3 - SUBSTITUIÇÃO PELO TERMO "EM DIREÇÃO"

Expressões locativas	Substituição
Olhar para um dos lados.	Em direção a um dos lados.

FONTE: A autora (2021).

- Os verbos “possui, tem, contém” podem ser substituídos pelo termo “com” que é uma preposição que foi classificada como topológica.

QUADRO 4 - SUBSTITUIÇÃO PELO TERMO "COM"

Expressões locativas	Substituição
A rua possui/tem/contém um aclave.	Rua com um aclave.
A praça possui/tem/contém muitas edificações históricas.	Praça com muitas edificações históricas.

FONTE: A autora (2021).

- Os verbos “há” e “existe” podem ser substituídos pelo verbo “ter” mas não podem ser substituídos pelo termo “com”. Estes verbos ficaram sem classificação.

4.1.5 Separação dos termos espaciais a partir dos mesmos *locata* e *relata*

De posse dos resultados das expressões locativas das nove tabelas que contém os *relata*, *locata* e os termos de relações espaciais, foi possível fazer uma separação dos termos que foram utilizados com os mesmos *locata* e *relata*. Estes termos deveriam ser considerados semanticamente similares porque estão sob uma mesma situação, numa mesma região e tem os mesmos elementos de *locata* e *relata*.

Os elementos utilizados como *locata* e *relata* foram agrupados quando citados por seus nomes próprios ou substantivos, quando da certeza de serem os mesmos elementos. Por exemplo, para a localidade BRA1 foram agrupados como mesmo *relata* e *locata* os termos: “cruzamento” e “esquina”, “rua” e “via”, “edificações históricas”, “prédio histórico” e “edifício de características coloniais”, “lugar” e “local”, “mercado” e “comércio”, entre outros. E os elementos implícitos foram agrupados de acordo com os elementos aos quais correspondiam, como: “acidente” e “implícito-acidente”. Desta maneira, aumentou-se a quantidade de pares de elementos e conseqüentemente isto possibilitou a análise de mais termos de relações espaciais quanto às suas similaridades. Todas as localidades foram trabalhadas da mesma maneira e com o mesmo raciocínio e a TABELA 7 ilustra

parte do resultado. A tabela completa e as demais tabelas para as outras localidades encontram-se na base de dados científicos (<http://dx.doi.org/10.5380/bdc/68>).

TABELA 7 - RELAÇÕES ESPACIAIS A PARTIR DOS MESMOS *LOCATA* E *RELATA* – TESTE ONLINE - BRA1

<i>Locata</i>	<i>Relata</i>	Termos de relações espaciais
acidente, implícito - acidente	cruzamento, esquina	(l) ocorreu em um (r) (l implícito-acidente) no (r) (l) aconteceu em um (r) (l) no (r) (l) na (r) (l) numa (r)
hotel	restaurante	(l) próximo ao (r) (l) na frente de (r)
muitas edificações históricas	praça	(r) possui (l) (l) ao redor (r)
prédio histórico de cor amarelo	praça	em frente a (r) ha (l) (l) em frente à (r)
muitas edificações históricas bem antigas	via	(r) tem (l) nessa (r) existem (l)
várias casas antigas	lugar	(r) possui (l) (r) com (l) entre (r plural) um (l)
construção	praça	anexo à (r) tem (l) na (r) existe (l) (l) adjacente a (r)
esquina	hotel	(l) próxima a (r) (l) perto do (r)
eu	praça	(l) estou localizado em uma (r)
praça	esquina	(l) estou situado perto de (r) (l) estou em frente (r) na (r) tem (l) neste (r) ha (l) (r) possui (l)

FONTE: A autora (2021).

Para a localidade BRA1 foram mencionados 18 grupos com mesmos *locata* e *relata*. Esta maneira de análise contribuiu para entender geograficamente o que as relações espaciais podem representar.

A maioria dos termos pode ser substituída uns pelos outros, porém são de classes diferentes e podem ser representados geograficamente de diferentes maneiras, apesar de terem sido mencionados no mesmo contexto e com os mesmos elementos. Os exemplos abaixo e a FIGURA 20 ilustram uma destas situações, onde o ponto de observação está representado pelo símbolo da estrela vermelha:

- 1) “Estou situado perto de uma praça.” O círculo verde mostra que o usuário pode estar em qualquer lugar a qualquer distância (não necessariamente na frente) da praça.
- 2) “Estou em frente à uma praça.” O círculo laranja mostra que o usuário pode estar em qualquer uma das quatro laterais da praça que podem ser consideradas como as frentes da praça.
- 3) “Estou localizado em uma praça.” O polígono azul ilustra que o usuário pode estar em qualquer lugar dentro dos limites da praça.

FIGURA 20 - ESPACIALIZAÇÃO DOS TERMOS DE RELAÇÕES ESPACIAIS – BRA1



FONTE: A autora (2021).

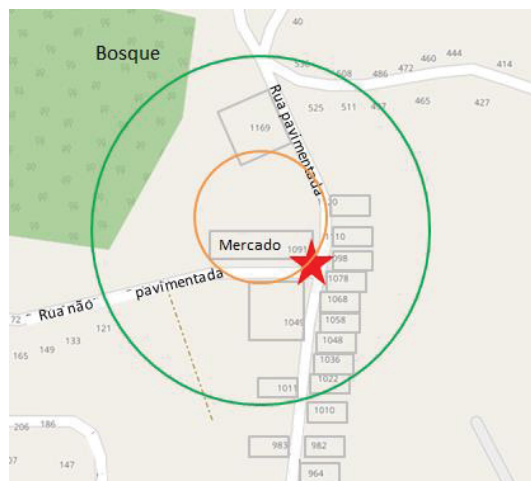
No caso de uma aplicação computacional, como o Projeto *Where am I?*, observando a situação geográfica da FIGURA 20 sugere-se que a interação do aplicativo seja realizada primeiramente a partir da escolha de termos de relações espaciais da classe “distância”. Esta classe, de acordo com a FIGURA 20, traria uma determinada área ou algumas determinadas áreas no mapa conforme ilustra o círculo na cor verde. A próxima iteração e escolha de termos de relações espaciais poderia ser da classe “projetiva”, que no caso acima estaria no círculo de cor laranja e traria uma área escolhida de uma determinada praça. E por fim, a escolha de termos da classe “topológica”, que no caso ilustrado, seria dentro dos limites da praça, ou seja, já seria o local desejado de localização do usuário. Desta maneira, a descrição partiria de termos de relações espaciais que representam ser mais distantes do ponto desejado até chegar ao local propriamente dito. Resumindo, a ordem para a interação do aplicativo referente às relações espaciais seria: termos da

classe “distância”, em seguida, termos da classe “projetiva” e por fim, termos da classe “topológica”.

A FIGURA 21 mostra mais um exemplo que a escolha dos termos pode seguir a ordem definida acima, agora para a localidade BRA2. O termo “próximo ao” é o mais distante, os termos “em frente ao” e “ao lado de” estão no mesmo nível de distância podendo ser em qualquer um dos lados do supermercado, que se encontra numa esquina; e são da classe projetiva. Percebe-se que para este exemplo não obtiveram termos para a classe topológica.

- 1) “Acidente próximo ao supermercado.” (círculo verde indica que o falante pode estar em qualquer lugar da linha verde)
- 2) “Acidente ao lado do supermercado.” e
- 3) “Acidente em frente ao supermercado.” (círculo na cor laranja mostra que estes dois termos estão sob a linha laranja).

FIGURA 21 - ESPACIALIZAÇÃO DOS TERMOS DE RELAÇÕES ESPACIAIS – BRA2



FONTE: A autora (2021).

Pode-se dizer que os termos não são semanticamente similares quando estão em diferentes classes e conseqüentemente a diferentes distâncias do *relatum*. Quando as descrições são espacializadas e transferidas para um mapa o termo “próximo” não significa a mesma coisa que “em frente” ou “ao lado” e sim “nas proximidades” do *relatum*. Já os termos “em frente” e “ao lado” (caso da FIGURA 21) apesar de serem de grupos diferentes são da mesma classe e apareceram sobre a mesma linha. Isto pode ocorrer em casos, por exemplo, de elementos como praças

e quadras de esporte, onde os elementos podem ser considerados com mais de uma frente e geralmente em esquinas ou quadras inteiras. Utilizando o egocentrismo e a linguagem natural percebe-se que as relações espaciais “em frente” e “ao lado” podem ser utilizadas com os mesmos elementos, por exemplo, para se encontrar um mercado em uma esquina não faria diferença.

As sequências de termos de relações espaciais que resultaram dos exemplos acima corroboram com o estudo de (FERREIRA; DELAZARI, 2019a) que explicam que no seu experimento, houve uma tendência da maioria dos usuários para o uso da relação “próximo” à medida que a distância entre os elementos referenciados aumenta.

4.2 COLETA DE DESCRIÇÕES DE LOCALIZAÇÃO - PRESENCIAL

Este teste consistiu em um experimento com um questionário *online* e duas tarefas com base na orientação espacial dos participantes. Teve como objetivo coletar descrições de localização em campo, ou seja, *in loco*, para a extração de termos de relações espaciais. Nas tarefas foram coletadas as descrições verbais em língua portuguesa brasileira e estas descrições proporcionaram as análises de quais foram os termos de relações espaciais mais recorrentes para ambientes *indoor*. As áreas para este teste foram pré-determinadas e os edifícios utilizados pertencem ao Campus III do Centro Politécnico da Universidade Federal do Paraná.

A partir das 66 descrições de localização coletadas, foi realizado o mesmo procedimento do primeiro teste, ou seja, o desmembramento em expressões locativas para que fosse possível realizar: 1) a extração e a contagem dos termos utilizados como relações espaciais e seus respectivos *relata* e *locata*, 2) o agrupamento e a classificação dos termos utilizados como relações espaciais, 3) a verificação da existência de novos termos utilizados como relações espaciais, e 4) a separação dos termos de relações espaciais a partir dos mesmos *locata* e *relata*.

4.2.1 Coleta das descrições de localização e desmembramento em expressões locativas

Os dados foram coletados de agosto a outubro de 2017. Participaram voluntariamente 33 indivíduos nativos da língua portuguesa brasileira, habituados a

se locomover pelo Campus III e familiarizados com as estruturas internas dos edifícios, o que facilitou para que os participantes dessem as descrições de localização. As 66 descrições verbais coletadas foram gravadas em 66 áudios, sendo 33 áudios com descrições de localização para cada uma das duas posições pré-determinadas. O experimento foi realizado individualmente para que não houvesse influência nas respostas dos participantes e foi gravado utilizando um aparelho celular.

Os voluntários eram alunos da graduação em Engenharia Cartográfica e de Agrimensura e da Pós-graduação de Ciências Geodésicas, de ambos os sexos, com idade de 19 a 42 anos. Desta maneira, considerou-se que diferenças entre o grau de escolaridade e as diferenças de faixa etária fossem minimizadas no experimento.

Após a coleta dos dados, iniciou-se a organização das descrições verbais para a realização da análise dos dados. As descrições verbais gravadas foram transcritas para o *Microsoft Word*. De posse das 66 descrições digitadas fielmente como foram descritas (<http://dx.doi.org/10.5380/bdc/69>), realizou-se o desmembramento em uma planilha Excel que resultou em 390 expressões locativas, sendo 203 para a Posição A e 187 para a Posição B. As frases que não eram locativas não foram consideradas.

As informações coletadas referentes às expressões locativas foram adicionadas em uma coluna do Excel para possibilitar o entendimento, análise e extração das relações espaciais. Ao término do desmembramento, a tabela continha as colunas referentes às descrições realizadas, as respectivas expressões locativas, os termos de relações espaciais utilizados entre o observador (*locatum*) e os elementos (*relata*), ou entre um elemento referenciado (*relatum*) e o elemento localizado (*locatum*), os *relata* e os *locata*.

4.2.2 Extração e contagem dos termos de relações espaciais e seus respectivos

locata e relata

A extração e a contagem dos termos de relações espaciais e seus respectivos *locata* e *relata* foi realizada do mesmo modo descrito no item 4.1.2.

A TABELA 8 mostra parte dos resultados da extração e da contagem dos termos de relações espaciais e seus respectivos *relata* e *locata* para a posição A. As

tabelas completas para as duas posições encontram-se na base de dados científicos (<http://dx.doi.org/10.5380/bdc/70>).

TABELA 8 - DESMEMBRAMENTO RELAÇÕES ESPACIAIS – POSIÇÃO A

Locata	Relata	Termos de relações espaciais
Porta	hidrante	(l) ao lado do (r)
implícito – eu	lixeiras	(r) ao fundo (l)
implícito – eu	edital	(r) a frente (l)
Eu	BNT	(l) to aqui no (r)
Eu	BNT	(l) to em frente ao (r)
Eu	cantina de alimentação	(l) to próximo da (r)
Eu	coordenação de mecânica	(l) to perto da (r)
Nós	departamento de mecânica	(l) estamos no (r)
Gabinetes	departamento de mecânica	(r) onde tem (l)
Nós	bloco 3	(l) estamos no (r)
PF1	bloco 3	(r) onde tem (l)
PF1	bloco 3	(r) um pouco à frente da onde tá (l)
PF4	bloco 3	(r) um pouco à frente da onde tá (l)
cantina	bloco 3	(r) um pouco à frente da onde tá (l)
sem locatum	coordenação do curso de mecânica	(r) está aqui
Eu	bloco de engenharia mecânica	(l) estou no (r)
implícito – eu	terceiro bloco do prédio das engenharias	(l) acredito que seja (r)
Eu	gabinete dos professores	(l) vejo (r)
coordenação do curso	implícito - eu	atrás (r implícito - eu) tem (l)
coordenação do curso	corredor	(l) no (r)
Cantina	início do corredor	(r) tem (l)
implícito – eu	elevador	(l) próximo do (r)
implícito - eu	cantina	(l) estou em frente a (r)
departamento de mecânica	implícito - eu	a (r implícito - eu) direita tem (l)
coordenação do curso de engenharia mecânica	implícito - eu	a (r implícito - eu) esquerda tem (l)
Gente	bloco de geologia	a frente (l) vai entrar no (r)
bloco da engenharia civil	cantina	atrás da (r) é o (l)

FONTE: A autora.

4.2.3 Agrupamento e a classificação dos termos de relações espaciais

Os termos de relações espaciais foram agrupados, da mesma maneira que o teste anterior, por gênero, plural, preposições, verbos com termos de relações espaciais, os próprios termos de relações espaciais e os verbos sem termos de relações espaciais. A TABELA 9 abaixo mostra o agrupamento para as duas posições A e B do termo de relação espacial “em frente”, que resultou em 18 grupos:

TABELA 9 - AGRUPAMENTO DO TERMO DE RELAÇÃO ESPACIAL “EM FRENTE” DO TESTE PRESENCIAL

Termo de relação espacial	Agrupamento do termo de relação espacial “em frente” - <i>INDOOR</i>
em frente	(l implícito - eu) estou em frente a (r), (l implícito - eu) estou em frente aos (r), (l implícito - eu) estou em frente da (r), (l implícito - eu) estou em frente ao (r), (l implícito - eu) estou na frente do (r)
	(l implícito - eu) em frente ao (r), (l implícito - eu) na frente da (r), (l implícito - eu) na frente do (r), (l implícito - eu) em frente a (r)
	(l) estou em frente a (r), (l) estou em frente ao (r), (l) estou na frente da (r), (l) estou na frente do (r)
	na (r implícito - eu) frente tem (l)
	(l) em frente a (r), (l) na frente da (r), (l) na frente do (r)
	(l) “to” em frente a (r), (l) “to” em frente ao (r), (l) “to” na frente do (r)
	na frente da (r) tem (l), na frente do (r) tem (l)
	(l implícito - eu) estou localizado em frente a (r)
	(l) bem na frente das (r), (l) bem na frente do (r)
	na frente (r) “vê” a (l), na frente (r) “vê” o (l)
	tem (l) na (r - implícito eu) frente
	(l implícito - eu) estou bem na frente da (r)
	(l) estou localizado em frente ao (r)
	(l) “to” aqui em frente ao (r)
	(l) “to” bem na frente do (r)
	na frente (r) tenho (l)
	na frente desta (r) tem (l) tem (l) na frente

FONTE: A autora.

A classificação dos termos de relações espaciais para este teste também foi realizada a partir das classes: topológica, direcional, projetiva, cardinal, distância, adjacência ou orientação.

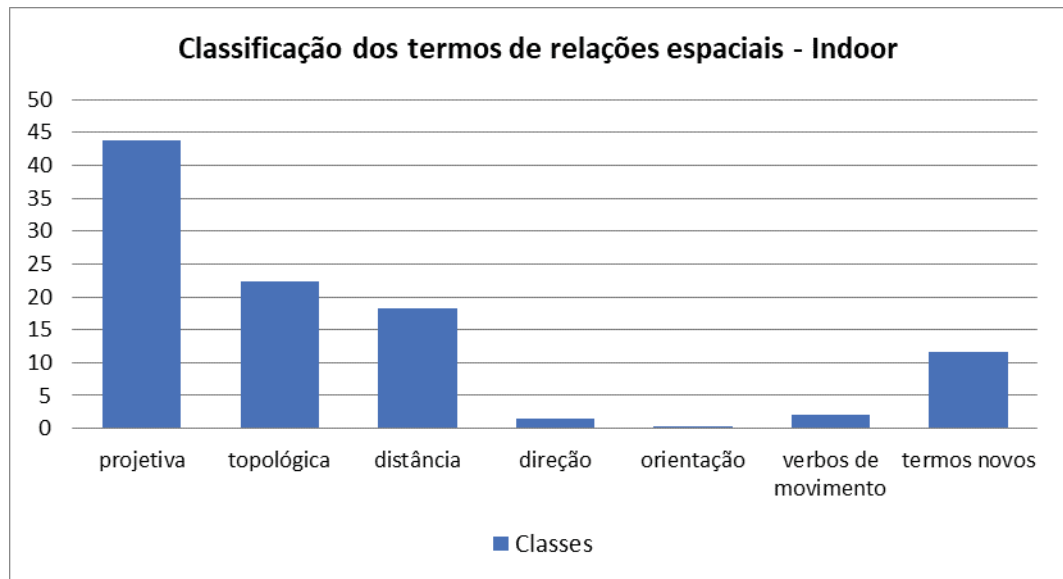
O resultado foi a TABELA 10 com os 390 termos de relações espaciais resultantes das Posições A e B, distribuídos em 41 grupos de relações espaciais com suas respectivas classificações.

TABELA 10 - AGRUPAMENTO E CLASSIFICAÇÃO DOS TERMOS DE RELAÇÕES ESPACIAIS DO TESTE PRESENCIAL

Classificação	Agrupamento dos termos de relações espaciais - Indoor	total
Projetiva	em frente, na frente	52
	a direita, a esquerda	26
	lado direito, lado esquerdo	20
	do lado, ao lado	17
	a frente	17
	atrás	9
	de frente	8
	depois	4
	antes	2
	ao fundo	2
	lado oposto	2
	pela frente	1
	ao fim	1
	trás	1
	nas costas	1
	outro lado	1
	um dos lados	1
	dois lados	1
lateral	1	
no final	1	
no final oposto	1	
abaixo	1	
acima	1	
Topológica	no, na	76
	entre	4
	dentro	2
	nesse	2
	em um, num	2
	neste	1
Distância	próximo	30
	perto	28
	aqui	11
	xxx metros	1
	daqui	1
Direção	para frente	2
	para baixo	2
	dá acesso	1
	desde... até...	1
Orientação	ortogonal	1
Verbos de movimento	vai, vem, passa, encontra, subi, interliga, leva, liga	8
Termos novos	verbos	45
	com	1
Total		390

FONTE: A autora (2021).

Para este teste, assim como para o primeiro teste, a classe projetiva também se sobressaiu sobre as demais classes, ficando com 44% das citações, conforme ilustra o GRÁFICO 2.

GRÁFICO 2 - CLASSIFICAÇÃO DOS TERMOS DE RELAÇÕES ESPACIAIS – *INDOOR*

FONTE: A Autora.

A TABELA 11 diz respeito ao agrupamento dos 45 verbos citados como relações espaciais que estão agrupados na TABELA 10.

TABELA 11 - AGRUPAMENTO DOS VERBOS DO TESTE PRESENCIAL

Agrupamento dos verbos	Total
(r) tem (l), (r implícito - local) tem (l), (r implícito - corredor) tem (l), (r) aonde tem (l), (r) consigo ver (l), (r) da para ver (l), (l implícito - corredor) da para ver (r), dá para ver (r), (l) vê o (r), (l) vejo (r)	22
(l) que da pra (r), (l) que dá para (r)	8
(l implícito - eu) estou olhando para (r), (r) olhar para (l), quando (l) olha para (r)	4
tem (l) que dá para o (r), tem (l) que dão para (r)	3
(l) acredito que seja (r)	2
(r implícito - lugar) aonde vai ter (l)	1
(r) onde exatamente (l) estou	1
(r) vai ver (l)	1
(l) pode ver (r)	1
(l) posso ver (r)	1
TOTAL	45

FONTE: A autora (2021).

4.2.4 Novos termos utilizados como relações espaciais

Para o contexto *indoor* também apareceu o termo “com” (é uma preposição que pode ser classificada como topológica), porém citado apenas uma vez, sendo mais expressiva a utilização de verbos para este contexto, conforme ilustrou a TABELA 11.

O verbo “ter” também se sobressaiu neste teste, tendo sido citado 22 vezes, assim como no contexto *outdoor*. Abaixo segue alguns exemplos de expressões locativas com os verbos citados que podem ser substituídos por termos de relações espaciais que permitem uma classificação:

- “consigo ver, vejo, vê” podem ser substituídos pelo termo “à frente” nas expressões locativas e serem inseridos na classe projetiva. Assim como ocorreu no teste *Outdoor*.

QUADRO 5 - SUBSTITUIÇÃO PELO TERMO "À FRENTE" - *INDOOR*

Expressões locativas	Substituição
Eu consigo ver a cantina.	À (minha) frente tem a cantina.
Eu vejo um gabinete de professores.	À (minha) frente tem um gabinete de professores.

FONTE: A autora (2021).

- “olhar para, que dá para/prá, que dão para, olhando para, que vai para,” podem ser substituídos por “em direção” e classificados como de “direção”. Assim como ocorreu no teste *Outdoor*.

QUADRO 6 - SUBSTITUIÇÃO PELO TERMO "EM DIREÇÃO" - *INDOOR*

Expressões locativas	Substituição
Se você olhar para um dos lados do corredor.	Em direção a um dos lados do corredor.
Passarela de vidro que dá para a biblioteca.	Passarela de vidro em direção à biblioteca.
Porta que dá para um jardim fechado.	Porta em direção a um jardim fechado.
Corredor que vai para as salas.	Corredor em direção às salas.

FONTE: A autora (2021).

- A expressão “que vem da ... pra ...” pode ser substituída por “de/da/do ... para ...” e classificado como de “direção”.

QUADRO 7 - SUBSTITUIÇÃO PELO TERMO "DE ... PARA..."

Expressões locativas	Substituição
O corredor que vem da biblioteca pra administração.	O corredor da biblioteca para a administração.

FONTE: A autora (2021).

- A expressão “dá para ver” pode ser substituída pelos verbos “tem/existe/há”, mas não se encaixam em outro grupo, portanto não tem como classificar.

QUADRO 8 - SUBSTITUIÇÃO PELOS TERMOS "TEM/EXISTE/HÁ"

Expressões locativas	Substituição
Dá para ver a biblioteca.	Tem/existe/há uma biblioteca.

FONTE: A autora (2021).

- A expressão “onde exatamente estou” pode ser substituída por “no, na”, podendo ser classificada como “topológica”.

QUADRO 9 - SUBSTITUIÇÃO PELOS TERMOS "NO/NA"

Expressões locativas	Substituição
Meio do corredor onde exatamente eu estou.”	Eu estou no meio do corredor.

FONTE: A autora (2021).

- As expressões “posso ver, pode ver” podem ser classificadas como projetivas e substituídas pelo termo “à frente”. Assim como ocorreu no teste *Outdoor*.

QUADRO 10 - SUBSTITUIÇÃO PELO TERMO "À FRENTE"

Expressões locativas	Substituição
Posso ver todo o bloco PI.	À minha frente está/tem o bloco PI.
Você pode ver a passarela.	À sua frente está/tem a passarela.

FONTE: A autora (2021).

Isto ocorre porque, ao se fazer a descrição em linguagem natural as pessoas têm total liberdade de escolher os termos a serem utilizados de acordo com seu vocabulário e sua percepção espacial (CAMBRIA; WHITE, 2014; CLARK, 1973; LANDAU; JACKENDOFF, 1993; MACHADO, 2001). As expressões que contém os verbos “ver, olhar, avistar” caracterizam orientação e fazem parte da percepção humana do espaço, concordando com Tenbrink, Bergmann e Konieczny (2011) que analisou a verbalização de rotas em ambientes *indoor*. Segundo a autora, os verbos de percepção e orientação, até certo ponto indicam falta de conhecimento do ambiente, enquanto as referências a entidades espaciais concretas transmitem segurança na informação espacial. Contudo, isto tem importância a partir do

momento que são conhecidos os termos da descrição, como seus pontos e relações espaciais. E ainda, concordando com a autora, de que dados como estes são úteis em uma variedade de aplicações práticas, como sistemas de diálogos e a melhoria de ambientes espaciais principalmente *indoor*.

4.2.5 Separação dos termos espaciais pelos mesmos *relata* e *locata*

A partir das tabelas com os resultados de *relata*, *locata* e relações espaciais das duas posições foi possível fazer a separação dos termos de relações espaciais que foram utilizados com os mesmos *locata* e *relata*, da mesma maneira que foi feita para as nove localidades do contexto *outdoor*.

Como os elementos foram separados por posição, para a Posição A agrupou-se como mesmo *relata* e *locata* os termos: cantina, cantina interna do prédio, cantina de alimentação, cantina interna dos blocos; as citações de maneiras diferenciadas para bloco, gabinete, coordenação e departamento. E os elementos implícitos foram agrupados de acordo com os elementos aos quais correspondiam, como por exemplo, “eu” e “implícito-eu”.

A TABELA 12 ilustra parte do resultado da Posição A, sendo que as tabelas completas para as duas posições se encontram na base de dados científicos (<http://dx.doi.org/10.5380/bdc/71>). Coincidentemente as duas posições tiveram 32 grupos com os mesmos *locata* e *relata*.

TABELA 12 - RELAÇÕES ESPACIAIS A PARTIR DOS MESMOS *RELATA* E *LOCATA* – TESTE PRESENCIAL - POSIÇÃO A (continua)

Locata	Relata	Termos de relações espaciais
bancos	corredor	no final oposto do (r) tem (l) (l) na lateral (r) (r) tem (l)
cantina	implícito - eu	a (r implícito - eu) frente tem (l) mais para (r implícito - eu) frente no lado esquerdo tem (l) na (r implícito - eu) frente tem (l) a (r implícito - eu) direita tem (l)
cantina	prédio da PF	(l) no (r) (l) dentro do (r)
coordenação do curso de engenharia mecânica	implícito - eu, eu	na (r implícito - eu) direita tem (l) a direita (r) tenho (l) do lado direito (r) olho a (l) a (r implícito - eu) esquerda tem (l) atrás (r implícito - eu) tem (l)

TABELA 12 - RELAÇÕES ESPACIAIS A PARTIR DOS MESMOS *RELATA* E *LOCATA* – TESTE PRESENCIAL - POSIÇÃO A (conclusão)

Locata	Relata	Termos de relações espaciais
coordenação do curso	corredor	(l) no (r) (l) no (r)
		(l) próximo a (r) (l) um pouco depois da (r) (l) depois da (r) (l) que dá de frente pra (r) (l) que dá pra (r) (l) acima da (r)
implícito - eu, eu	cantina	(l implícito - eu) estou próximo da (r) (l implícito - eu) estou próximo a (r) (l) consigo ver (r) daqui (l implícito - eu) próximo das (r) (l implícito - eu) estou próximo a (r) (l implícito - eu) estou localizado em frente a (r) (l implícito - eu) perto da (r) (l implícito - eu) estou próximo a (r)

FONTE: A autora (2021).

4.3 COMPARAÇÃO E VERIFICAÇÃO DAS HIPÓTESES PROPOSTAS PARA OS TESTES *ONLINE* E PRESENCIAL

4.3.1 Comparação dos termos de relações espaciais nos dois testes

Os testes nos contextos *outdoor* e *indoor* foram realizados com a finalidade principal de se obter um grupo de termos de relações espaciais mais utilizados na linguagem natural do português brasileiro e também de responder às duas hipóteses propostas quando da criação dos testes. Os números gerais da coleta dos dados encontram-se na TABELA 13, onde o número total de palavras (*word tokens*) e a quantidade de tipos de palavras (*word types*) foi calculado pelo programa livre *AntConc* versão 3.5.8.

TABELA 13 – NÚMEROS GERAIS DA COLETA DE DADOS

Coleta de dados	<i>Outdoor</i>	<i>Indoor</i>
Descrições coletadas	121	66
Número total de palavras (<i>word tokens</i>)	4883	4072
Quantidade de tipos de palavras (<i>word types</i>)	907	527
Relações espaciais	557	390
Agrupamentos	59	41

FONTE: A autora (2021).

O agrupamento dos termos de relações espaciais das nove localidades no contexto *outdoor* (TABELA 4) resultou em 59 grupos de termos; e o agrupamento no contexto *indoor* (TABELA 10) resultou em 41 grupos.

Pode-se analisar que apesar das descrições de localização terem sido realizadas para ambientes com diferentes configurações e diferentes contextos, os termos de relações espaciais não sofreram grande variação na utilização e na grande maioria não foram específicos para cada tipo de ambiente. Apesar das descrições serem em linguagem natural e os participantes terem a liberdade de as descreverem como quiserem, os termos de relações espaciais foram citados de uma maneira geral sem variar muito em preferência de utilização.

A comparação dos resultados encontrados nos dois testes possibilitou a verificação dos termos de relações espaciais mais utilizados nos dois contextos apresentados da língua portuguesa brasileira. Esta comparação resultou em 79 grupos de termos, conforme TABELA 14. E, desta maneira, foi possível realizar a extração dos termos mais mencionados para serem utilizados no próximo teste.

TABELA 14 - TERMOS DE RELAÇÕES ESPACIAIS *OUTDOOR* E *INDOOR*

(continua)

Termos de relações espaciais <i>outdoor</i> e <i>indoor</i>	total	Termos de relações espaciais <i>outdoor</i> e <i>indoor</i>	total	
em frente, na frente	148	lado de dentro, lado de fora lado oposto na outra nas proximidades nesse para baixo para frente perpendicular sob x metros acima do	2	
no, na	132			
verbos	98			
próximo	71			
em um, em uma, em que, numa, num	55			
com	44			
perto do, perto de	42			
ao lado de, do lado da	39			
lado direito, lado esquerdo	33			
direita, esquerda	32			
à frente	23			
do outro lado da	13			
Entre	12			
Aqui				
de frente	11	acesso ao adjacente anexo ao fim ao final cujo fim cujo fundo dá acesso daqui dentre	1	
atrás da	10			
em um lado, de um lado, neste, nesta	8			
depois				
em direção a, em direção ao	5			
x distância				
Abaixo	4			desde... até... dois lados frente a frente
lado par, lado ímpar				

TABELA 14 - TERMOS DE RELAÇÕES ESPACIAIS *OUTDOOR* E *INDOOR*

(conclusão)

Termos de relações espaciais <i>outdoor</i> e <i>indoor</i>	total	Termos de relações espaciais <i>outdoor</i> e <i>indoor</i>	total
Antes ao fundo ao sul, ao norte Após de frente em ambos os lados, em ambas as laterais em cima junto a por cima Sobre	3	inferior lado contrário lateral nas costas no final no final oposto no sentido oposta ortogonal outro lado para o lado pela frente trás um dos lados um lado da vai do xx para o yy	1
Adiante ao longo ao redor dentro direção norte, direção sul em	2	TOTAL	911

FONTE: A autora (2021).

Para o terceiro teste (associação livre) foram utilizados os termos citados mais do que 1% (10 vezes) do total da TABELA 14, com exceção do termo “aqui” que apareceu somente no contexto *indoor*, ou seja, o elemento estava situado nas proximidades do observador/falante (*locatum*) que funcionou ao mesmo tempo como entidade de referência (*relatum*) (BLÜHDORN, 2001), portanto este termo se baseia na distância e foi utilizado porque as pessoas estavam *in loco* e poderiam apontar para os elementos ou para as localizações e utilizá-lo (TEIXEIRA, 2005). E também com exceção dos termos “direita” e “esquerda” e “lado direito” e “lado esquerdo”, porque não foi fixada a direção para onde os voluntários deveriam ficar virados que possibilitassem distinguir estes dois lados. Sendo assim, foram utilizados os doze termos “EM FRENTE, NA, PRÓXIMO, EM, COM, PERTO, AO LADO, À FRENTE, DO OUTRO LADO, ENTRE, DE FRENTE e ATRÁS” e o termo “QUE TEM” que representou o verbo “ter” que foi o mais citado do agrupamento dos verbos utilizados como relação espacial. Isto porque o agrupamento dos verbos aparece entre os mais utilizados; e os termos que continham o verbo “ter” somaram 32 citações, então se o agrupamento dos verbos fosse considerado desmembrado na TABELA 14 de termos total, ele também entraria na seleção com mais de 10 citações. E o termo “HÁ” do agrupamento dos verbos que apareceu em cinco citações.

Uma análise mais detalhada do uso de verbos como relações espaciais utilizados nas descrições, tanto em combinação com mais vocábulos, quanto sozinhos, revela que o verbo “ter” foi utilizado no sentido “existir” e de “possuir”. Embora os verbos possessivos não sejam normalmente utilizados para expressar relações espaciais, os resultados mostraram que eles se sobressaem sobre os demais verbos. Estes verbos podem ser classificados como topológicos, no sentido de indicar que algum objeto está contido em outro objeto. Exemplos de expressões locativas com o verbo “ter” e “possuir”: 1) BRA1 – “A via tem muitas edificações históricas.” 2) NZ2 – “A ponte possui grades metálicas em ambos os lados.” Em Inglês foram utilizados os termos “*with, on, in, to have, to contain*” como variantes do verbo “ter” do português brasileiro (FAGUNDES et al., NO PRELO)

A indicação de termos de relações espaciais semanticamente similares para os verbos utilizados corrobora com o trabalho de Belouaer e Claramunt (2013) que fez a substituição de verbos de movimento por termos de relações espaciais no estudo de caso no qual modelou o conhecimento espacial de descrições de rota. Além do que combinações semânticas de verbos e preposições estão presentes na linguagem natural porque expressam movimento em um determinado percurso ou local, por isto devem ser levadas em consideração, de outro modo significaria negar a criatividade e a complexidade linguística da mente humana (KEWITZ, 2011).

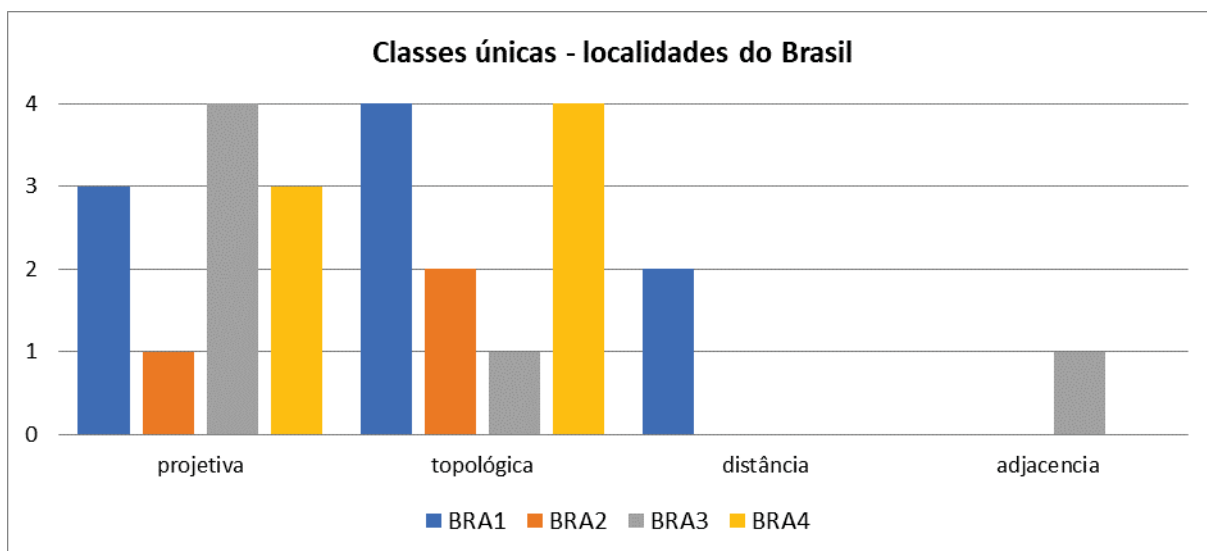
4.3.2 Comparação das classificações dos termos de relações espaciais encontrados com mesmos *relata* e *locata* dos dois testes

Conforme ilustrado no item 4.1.5 quando os mesmos *relata* e *locata* foram separados para a verificação de quais termos de relações espaciais foram utilizados, surgiu a necessidade da análise e da classificação destes termos para que os mesmos possam ser utilizados no projeto *Where am I?*.

Além de ter grupos com classificações únicas, como ilustram os GRÁFICOS 3 para as localidades no Brasil e 4 para as localidades da Nova Zelândia; existem os grupos com classificações conjuntas, conforme mostram os GRÁFICOS 5 e 6. As localidades NZ2 e NZ3 não apresentaram termos nas classes únicas.

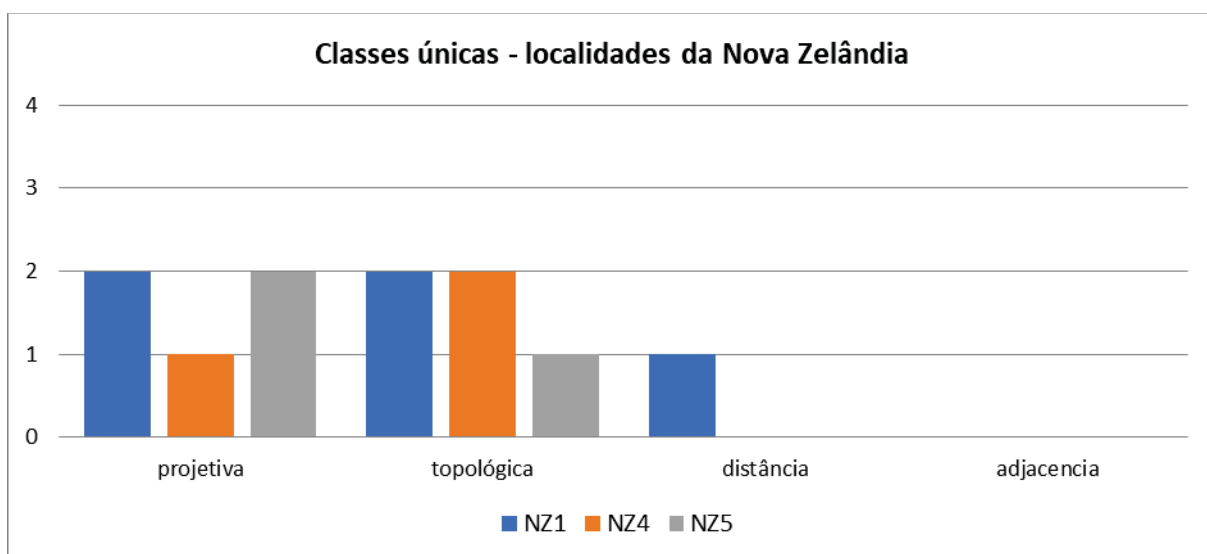
Percebe-se que para os dois grupos dos GRÁFICOS 3 e 4 a classe projetiva e a topológica têm maior destaque.

GRÁFICO 3 - CLASSIFICAÇÕES ÚNICAS – LOCALIDADES DO BRASIL



FONTE: A autora (2021).

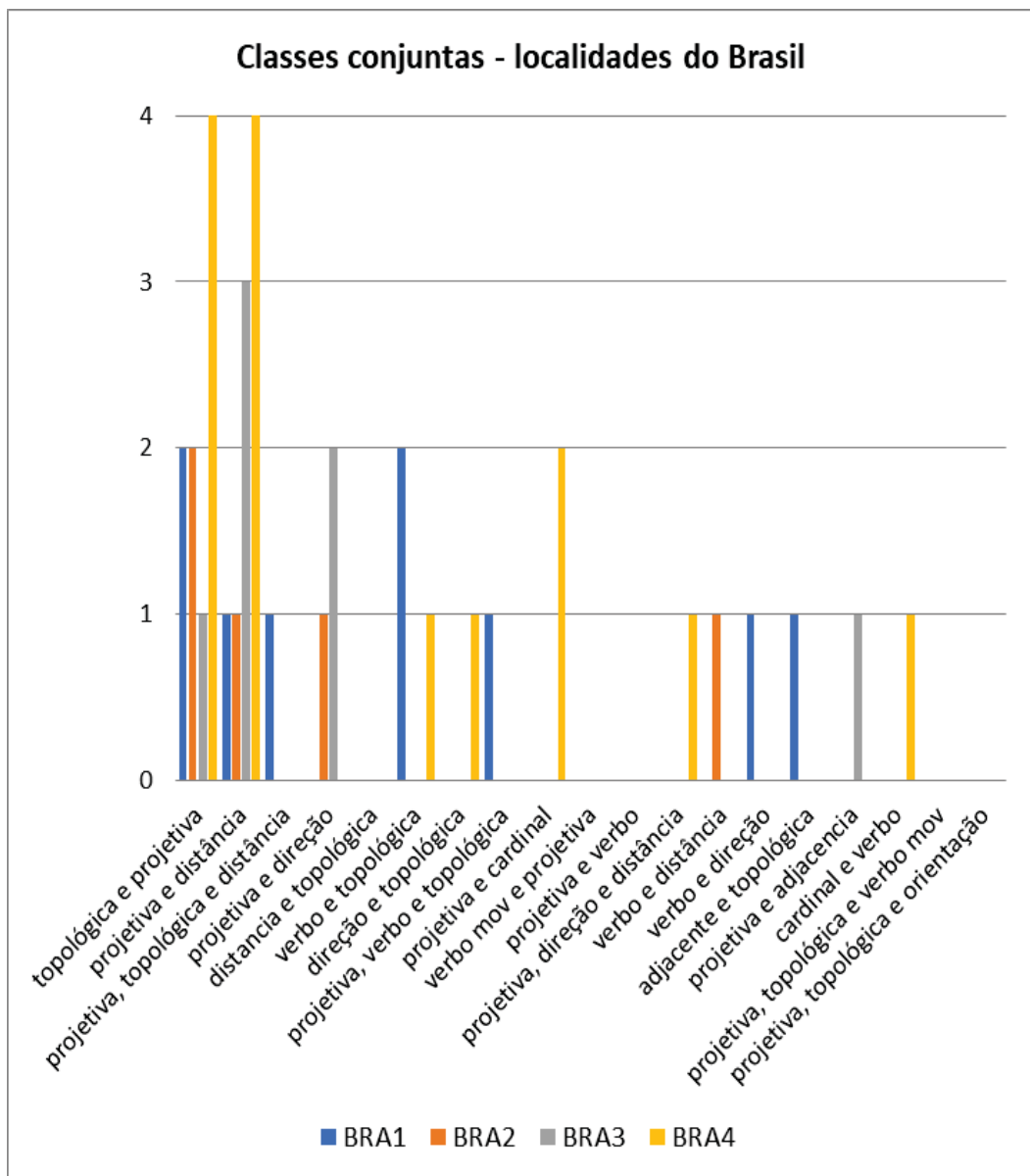
GRÁFICO 4 - CLASSIFICAÇÕES ÚNICAS - LOCALIDADES DA NOVA ZELÂNDIA



FONTE: A autora (2021).

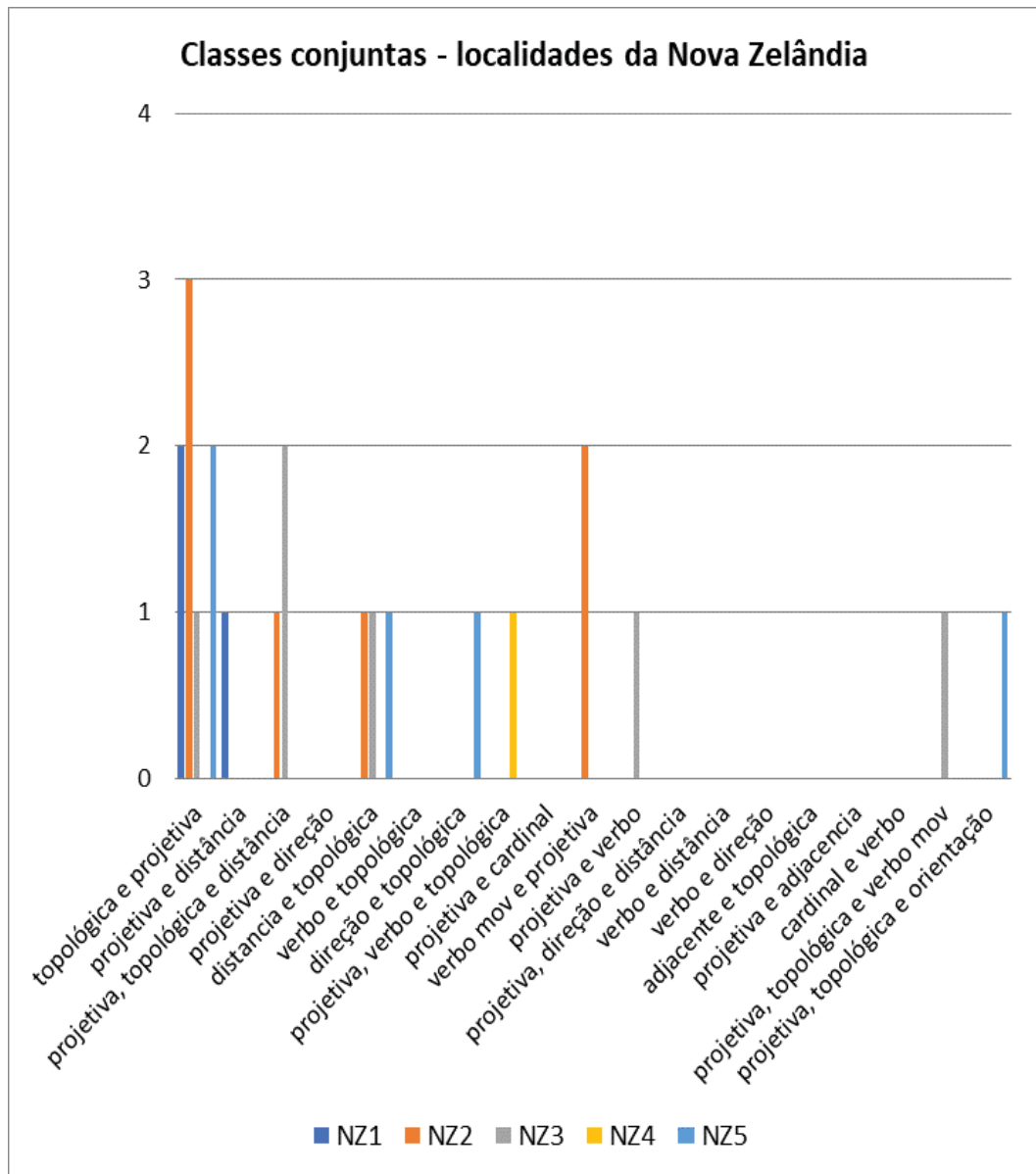
Os GRÁFICOS 5 e 6 mostram que as classes que foram unidas têm maior número, quando unidas à classe projetiva.

GRÁFICO 5 - CLASSIFICAÇÕES CONJUNTAS - LOCALIDADES DO BRASIL



FONTE: A autora (2021).

GRÁFICO 6 - CLASSIFICAÇÕES CONJUNTAS - LOCALIDADES DA NOVA ZELÂNDIA

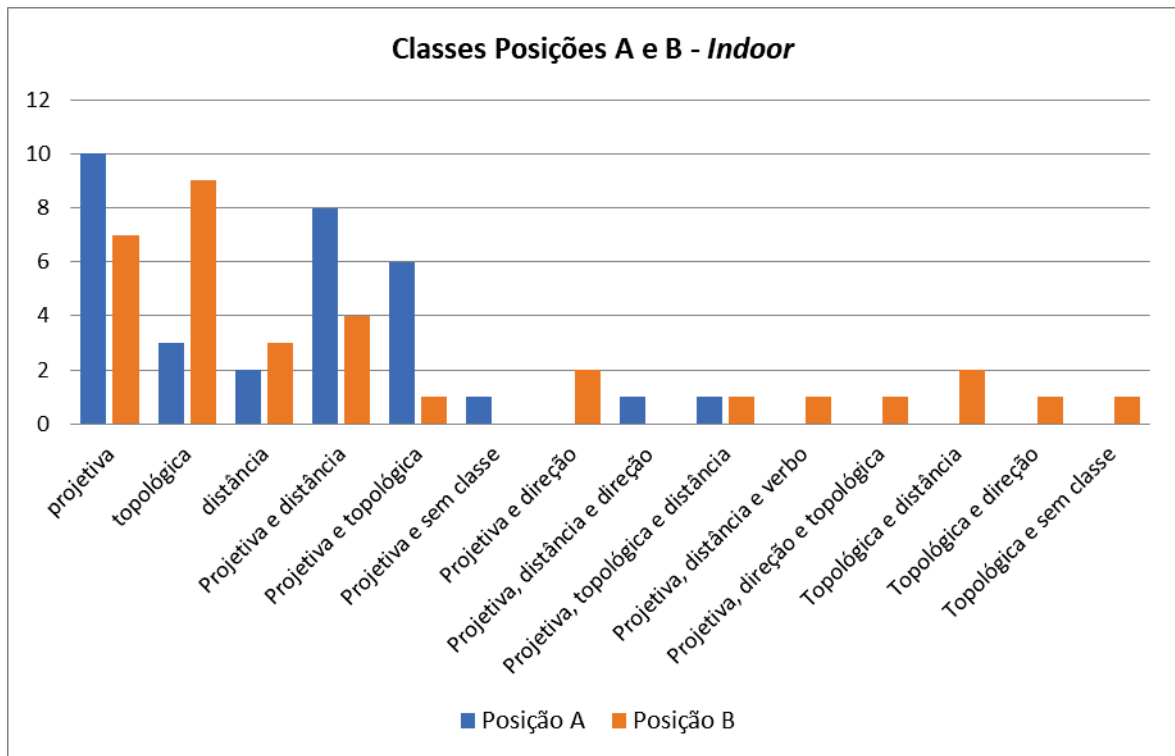


FONTE: A autora (2021).

O GRÁFICO 7 exemplifica como foi a aplicação da classificação para os termos de relações espaciais das duas posições do contexto *Indoor*. Pode-se observar que seis grupos de classificações foram utilizados nas duas posições. Sendo que a classe única projetiva acumula um total de 27%, seguida da topológica com 19% e projetiva/distância também com 19%.

O resultado da comparação das classes dos termos de relações espaciais confirma que a classe projetiva de relações espaciais é a classe mais presente na língua portuguesa brasileira.

GRÁFICO 7 - CLASSIFICAÇÃO DOS TERMOS DE RELAÇÕES ESPACIAIS DOS MESMOS LOCATA E RELATA – INDOOR



FONTE: A autora (2021).

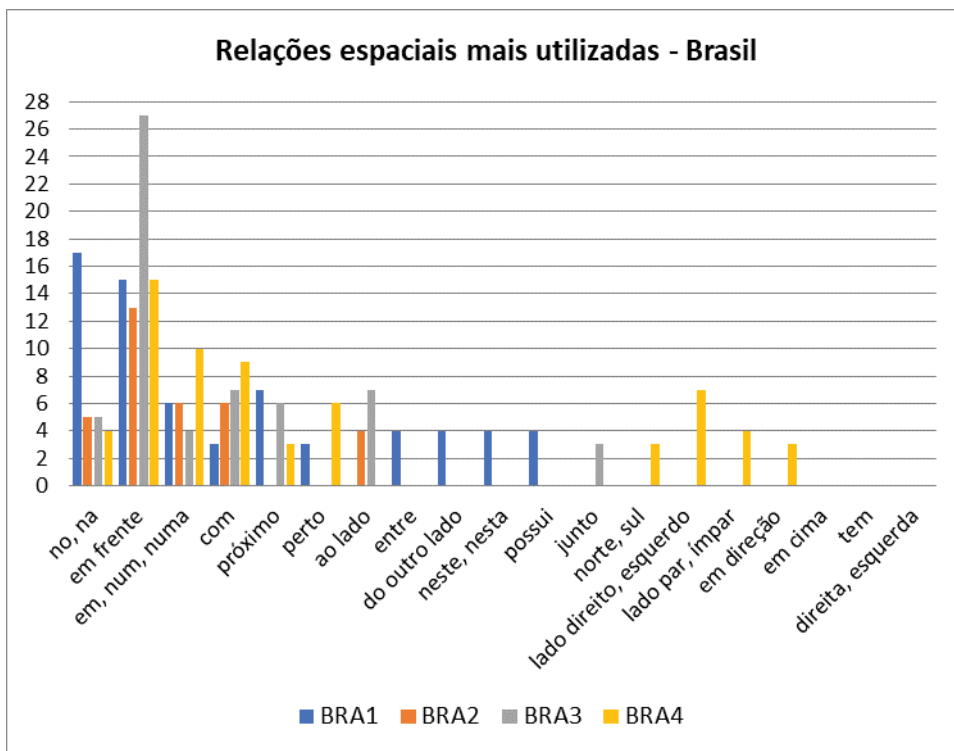
4.3.3 Verificação das hipóteses propostas para os dois testes

Os resultados e análises das hipóteses seguem a sequência listada na metodologia.

1ª hipótese para os testes *online* e em campo: Se a coleta das descrições de localização em linguagem natural se deu em dois contextos diferentes – *outdoor* e *indoor* - e em situações geográficas com diferentes configurações, então, pelo grau de liberdade que a linguagem natural oferece, os termos de relações espaciais que se sobressaem são diferentes para cada tipo de ambiente.

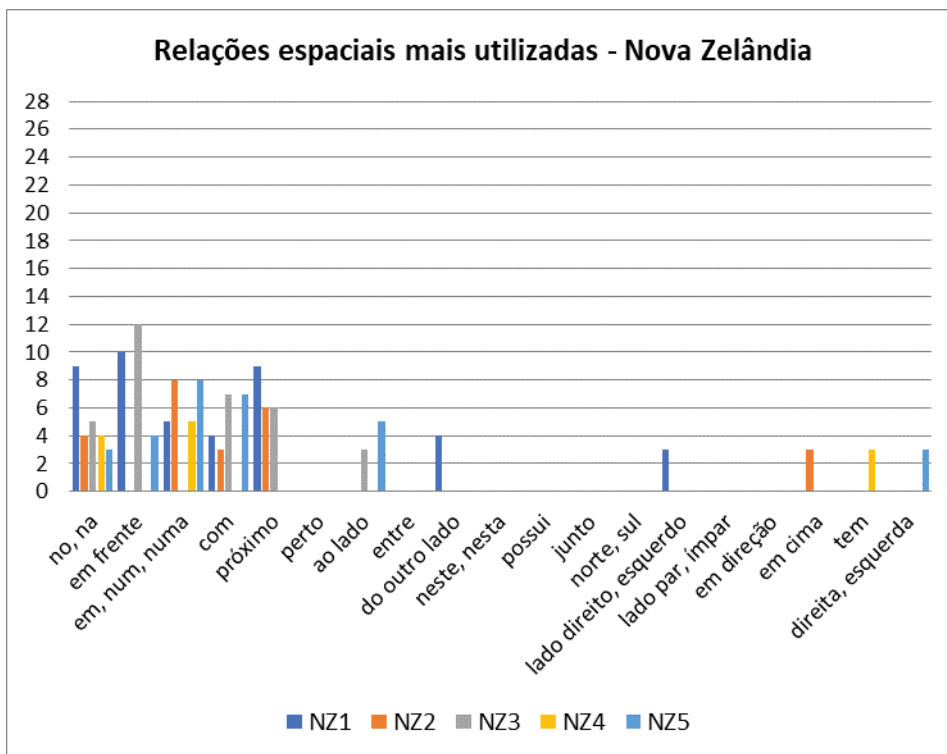
Resgatando os termos de relações espaciais, para cada uma das nove localidades, que foram citados mais do que três vezes, tem-se os GRÁFICOS 8 e 9 abaixo.

GRÁFICO 8 - RELAÇÕES ESPACIAIS MAIS UTILIZADAS NAS ÁREAS DO BRASIL



FONTE: A autora (2021).

GRÁFICO 9 - RELAÇÕES ESPACIAIS MAIS UTILIZADAS NAS ÁREAS DA NOVA ZELÂNDIA



FONTE: A autora (2021).

Percebe-se que os termos “no, na” aparecem nas nove localidades seguidos dos termos “com, em frente, em, num, numa, próximo” como os mais mencionados. Destacando-se o “com” que foi muito utilizado como relação espacial sendo mencionado mais que três vezes em oito localidades e apareceu em quarto lugar no total de termos do contexto *outdoor* no primeiro teste (TABELA 4). Isto comprova que esta preposição é muito utilizada para localização espacial.

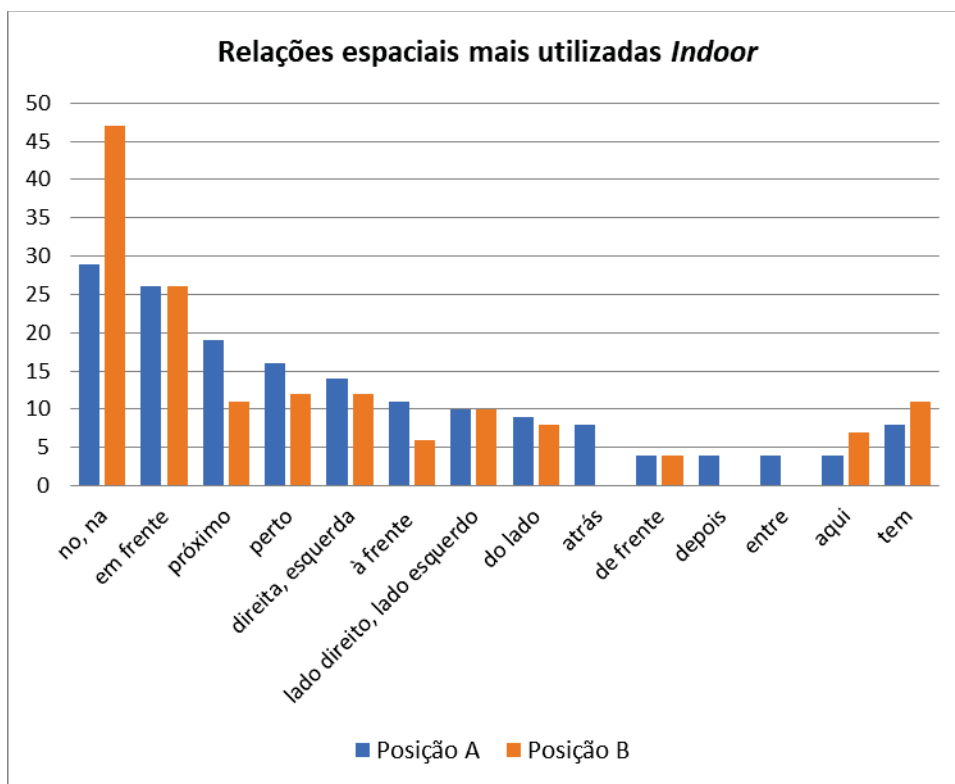
A GRÁFICO 10 apresenta as relações espaciais mais citadas no ambiente *outdoor*. Os resultados mostraram que foram mencionados 19 tipos de termos de relações espaciais, sendo que seis termos foram citados mais do que 10 vezes.

GRÁFICO 10 - RELAÇÕES ESPACIAIS MAIS UTILIZADAS NO CONTEXTO *OUTDOOR*



FONTE: A autora (2021).

No ambiente *indoor* 11 grupos de termos de relações espaciais foram mencionados em ambas as posições (GRÁFICO 11).

GRÁFICO 11 - RELAÇÕES ESPACIAIS MAIS UTILIZADAS NAS DUAS POSIÇÕES DO *INDOOR*

FONTE: A autora (2021).

O GRÁFICO 12 apresenta as relações espaciais mais citadas no ambiente *indoor*. Pode-se perceber que 10 grupos foram mencionados mais que 10 vezes (GRÁFICO 8), sendo que os termos “no, na, em frente, próximo e ao lado” também estão entre os mais utilizados no contexto *outdoor*.

GRÁFICO 12 - RELAÇÕES ESPACIAIS MAIS UTILIZADAS NO CONTEXTO *INDOOR*

FONTE: A autora (2021).

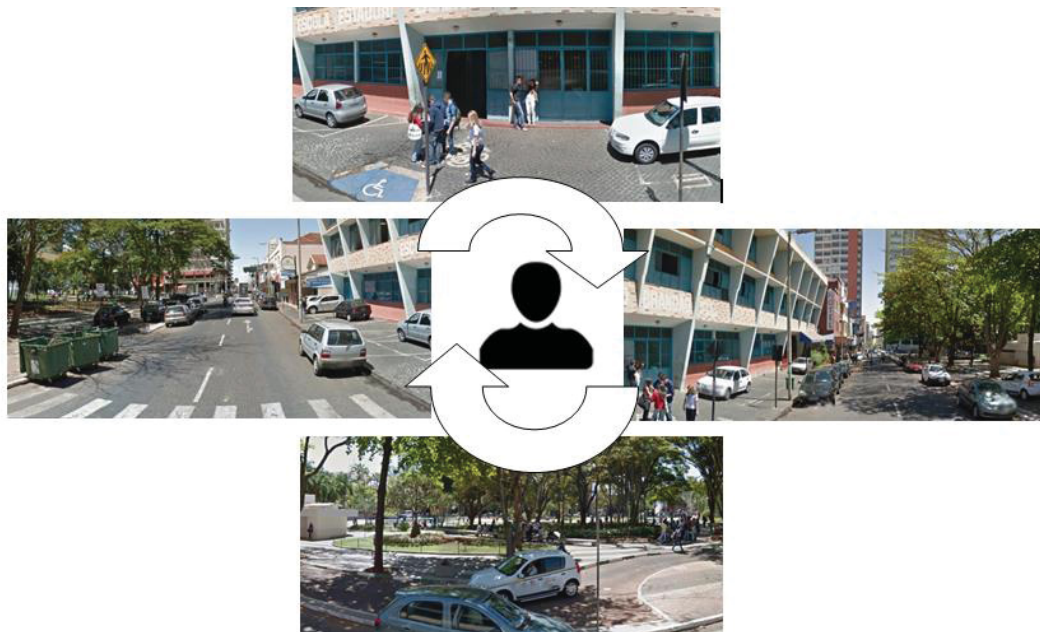
Com isto, afirma-se que na língua portuguesa brasileira existe um grupo de termos de relações espaciais que são mais utilizados e que variam pouco em ordem de preferência para descrições de ambientes com diferentes configurações e contextos. Ou seja, na língua portuguesa brasileira os termos: “em frente de/da/do, no/na/em, próximo, ao lado de, lado direito/esquerdo e perto” aparecerão em descrições de ambientes, sejam elas, com diferentes configurações e contextos variando somente a preferência de utilização. Isto porque 86% dos termos do contexto *outdoor* e 92% do contexto *indoor* foram mencionados mais do que 10 vezes nas 9 localidades e nas duas posições respectivamente. O restante mostra que os outros termos foram utilizados poucas vezes pelos voluntários. Portanto a hipótese é falsa, provando que para a língua portuguesa brasileira os termos mais utilizados não variam de acordo com o contexto.

2ª hipótese para os testes *online* e em campo: Se as pessoas fazem descrições de localização em linguagem natural em um determinado contexto e ambiente, de acordo com suas percepções do ambiente e do

espaço e ficando livres para optar pelos elementos a serem relacionados, então elas utilizarão os termos de relações espaciais no referencial centrado nos elementos, ou seja, as relações espaciais se darão entre os elementos existentes no ambiente.

No primeiro teste, no contexto *outdoor*, na maioria das vezes os voluntários utilizaram o eixo de referência centrado nos elementos, como por exemplo: “A escola está **entre** um hotel e um bar/restaurante” ou “Existe uma praça em frente à escola.”. A FIGURA 22 que corresponde à localidade BRA3 ilustra o ponto de vista do observador e os elementos que poderiam ser utilizados como *relata* e *locata* no caso do eixo centrado fora do observador.

FIGURA 22 - PONTO DE VISTA - LOCALIDADE BRA3



FONTE: A autora (2021).

Foram 450 expressões locativas utilizando o eixo centrado nos elementos das localidades descritas, mostrando que 86% dos voluntários preferiram este eixo. Contra 71 expressões locativas de modo egocêntrico (centrado no indivíduo), conforme ilustra a TABELA 15. Sendo que, destas 71 frases o sujeito “eu” apareceu em 18 frases e o termo “implícito – eu” apareceu em 57 frases, justificando a utilização de sujeitos ocultos conforme normas da língua portuguesa brasileira, como por exemplo: “Estou em uma rua sem saída.”

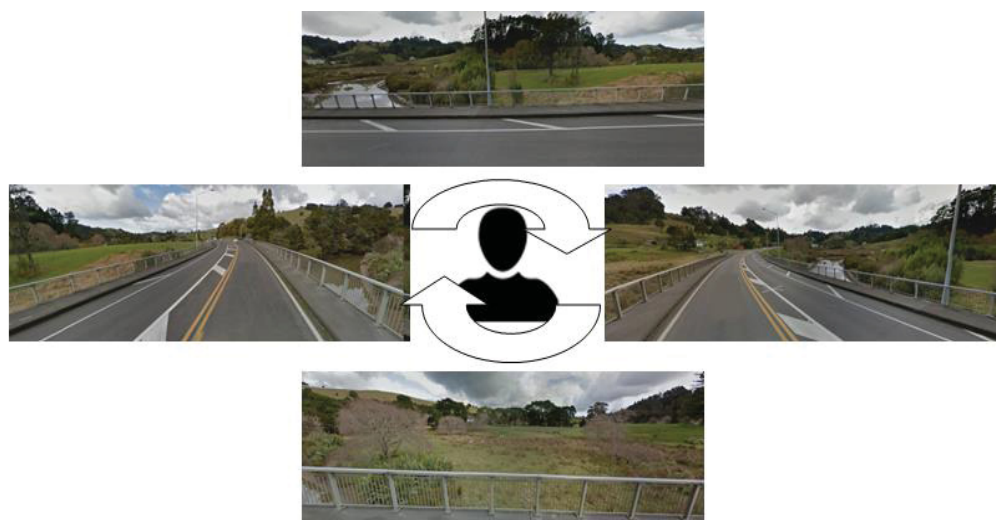
TABELA 15 – EIXO DE REFERÊNCIA - *OUTDOOR*

Localidade	Centrado no indivíduo (%)	Centrado no elemento (%)	Total de termos citados
BRA1	14	86	88
BRA2	17	83	52
BRA3	17	83	77
BRA4	19	81	85
NZ1	7	93	56
NZ2	9	91	54
NZ3	8	92	48
NZ4	15	85	20
NZ5	12	88	41
TOTAL	14%	86%	521

FONTE: A autora (2021).

A principal justificativa é que o teste foi realizado *online*, portanto os voluntários se preocuparam em procurar os dois elementos, o *relatum* e o *locatum* para fazer a localização. Apesar de o enunciado *outdoor* pedir para que os voluntários se imaginassem testemunhando um acidente, ou seja, deveriam ter se colocado nesta situação e em posição egocêntrica. Outra possível justificativa seria que os observadores escolheram o eixo centrado nos elementos pela quantidade de elementos disponíveis nas localidades. Mesmo a localidade NZ2 que é rural, obteve 90% das frases utilizando o eixo centrado nos elementos (FIGURA 23).

FIGURA 23 - PONTO DE VISTA - LOCALIDADE NZ2



FONTE: A autora (2021).

No segundo teste, no contexto *indoor*, houve 231 expressões locativas no modo egocêntrico, equivalendo a 59,23% do total de 390 expressões analisadas, de acordo com a TABELA 16. Este resultado suporta a justificativa do egocentrismo ser

maior quando as pessoas estão *in loco*, mostrando a dificuldade dos voluntários em se imaginar nos cenários *online* mesmo que o enunciado peça.

TABELA 16 - EGOCENTRISMO TESTE *INDOOR*

Posições	Centrado no indivíduo (%)	Centrado no elemento (%)	Total de termos citados
A	65	35	203
B	53	47	187
TOTAL	59%	41%	390

FONTE: A autora (2021).

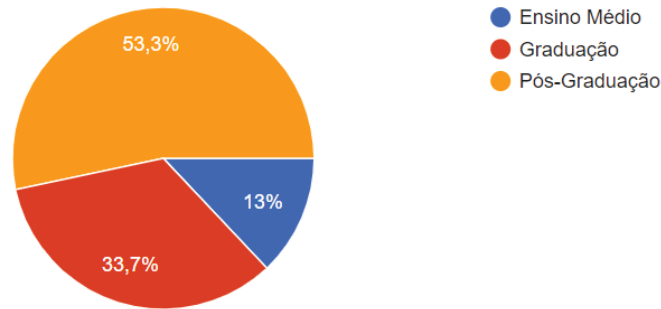
Portanto a hipótese se verifica no primeiro teste e não se verifica no segundo teste. Ficando como sugestão testes com as situações invertidas, ou seja, fotografias das quatro posições em ambiente *indoor* para as descrições de localização e *in loco* para ambientes *outdoor*.

4.4 TESTE DE ASSOCIAÇÃO LIVRE DE RELAÇÕES ESPACIAIS - SIMILARIDADE

Este teste foi realizado em outubro e novembro de 2019. Foram coletadas 270 respostas que totalizaram 3780 termos de relações espaciais que foram agrupados pelos 14 termos das frases que continham os termos mais citados nos dois testes de coleta de dados realizados anteriormente, na linguagem natural do português brasileiro: EM FRENTE, NA, QUE TEM, EM, PRÓXIMO, COM, PERTO, AO LADO, À FRENTE, DO OUTRO LADO, ATRÁS, ENTRE, DE FRENTE e HÁ.

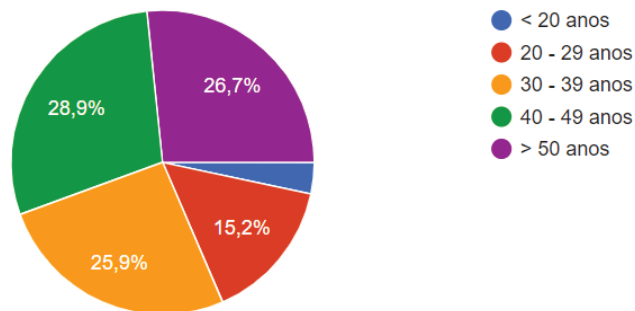
Foram excluídas 3 respostas onde dois dos participantes responderam que sua língua nativa era a língua Espanhola e um participante tinha o Ucrâniano como língua nativa. Portanto, foram utilizadas 267 respostas, sendo que mais da metade dos voluntários foram da pós-graduação (FIGURA 24) e a grande maioria tinha acima de 30 anos de idade (FIGURA 25). As respostas vieram de participantes das áreas técnica, biológicas e humanas. A participação do gênero feminino superou bastante a do gênero masculino, ficando com 76,3%.

FIGURA 24 – ESCOLARIDADE DOS PARTICIPANTES



FONTE: A autora (2021).

FIGURA 25 – FAIXA ETÁRIA DOS PARTICIPANTES



FONTE: A autora (2021).

Este teste serviu para confirmar os sinônimos de relações espaciais utilizados pelos voluntários. Estes sinônimos foram usados para a formação dos grupos semanticamente similares. O teste foi realizado sem a ilustração das situações, e todas as respostas foram obrigatórias, a fim de ter o mesmo número total de respostas para as 14 frases, que se encontram listadas abaixo:

1. Eu estou EM FRENTE À rodoviária.
2. Eu estou NA esquina de duas ruas movimentadas.
3. HÁ muitos prédios comerciais na rua em que estou.
4. Eu estou localizado em uma região QUE TEM vários prédios históricos.
5. Eu estou EM UM parque.
6. Eu estou PRÓXIMO DE um hospital.
7. Eu estou em um cruzamento COM quatro ruas.
8. Eu estou PERTO DO cemitério.
9. Eu estou AO LADO DE um supermercado.
10. À FRENTE de onde estou situado está o Centro Histórico da cidade.

11. DO OUTRO LADO DA rua em que estou tem uma universidade.

12. A maior ponte da cidade está ATRÁS DE MIM.

13. Eu estou ENTRE a praça e a prefeitura.

14. Eu estou DE FRENTE AO Palácio do Governo.

A intenção do teste não foi de testar a velocidade das respostas com contagem de tempo para responder e não foi aplicada qualquer situação de estresse (FERREIRA, 2019), para que os usuários ficassem livres para responder. Neste teste não houve respostas certas ou erradas, todas as respostas foram analisadas, por se tratar de linguagem natural.

O primeiro passo foi a exportação dos dados para uma planilha Excel onde as frases foram analisadas, contadas e agrupadas preliminarmente, resultando em uma tabela para cada um dos 14 termos principais das frases acima.

Na TABELA 17 seguem os resultados do agrupamento inicial para a relação espacial “em frente à” referente à primeira frase (35 grupos de termos). Observa-se que o total de respostas é 273 porque dois voluntários escreveram dois termos no espaço em branco e outros dois voluntários escreveram três termos no espaço em branco. Portanto, esta tabela totaliza 267 + 6 termos escritos a mais. Porém, os termos: “não, entrada principal, em casa, no acostamento da rodovia, no portão da” foram desconsiderados porque os voluntários alteraram o *relatum* da frase, e não utilizaram termos de relações espaciais. As tabelas contendo os agrupamentos iniciais dos 14 termos encontram-se na base de dados científicos (<http://dx.doi.org/10.5380/bdc/72>).

TABELA 17 - AGRUPAMENTO INICIAL DOS TERMOS DE RELAÇÕES ESPACIAIS - “EM FRENTE”
(continua)

Em frente à	Total
na frente, na frente da	67
diante, diante da/a/à	50
de frente, de frente a/para, defrente, defrente a/à/da	45
na	33
de frente, de frente a/à/com/com a/para/para a	13
próximo	13
em frente à/da/a	5
perto da	4
perante	4
a frente, a frente da	3
adiante, adiante da	3
ao lado, do lado	3

A quantidade elevada de termos atribuídos aos termos principais de relações espaciais das frases se justifica porque de acordo com Cambria; White (2014) os seres humanos como processadores de texto não tem limitações, pois, cada palavra que lhes vem à mente ativa uma cascata de conceitos semanticamente relacionados, episódios relevantes e expediências sensoriais. Isto lhes permite a conclusão de tarefas complexas de maneira rápida e sem esforço, como desambiguação das palavras, vinculação textual e rotulação semântica, o que não é o caso no PLN. Por isto, para este estudo foram elaboradas regras que filtrassem os termos para um resultado mais objetivo com relação à descoberta das relações espaciais semanticamente similares que possam vir a ser utilizadas computacionalmente.

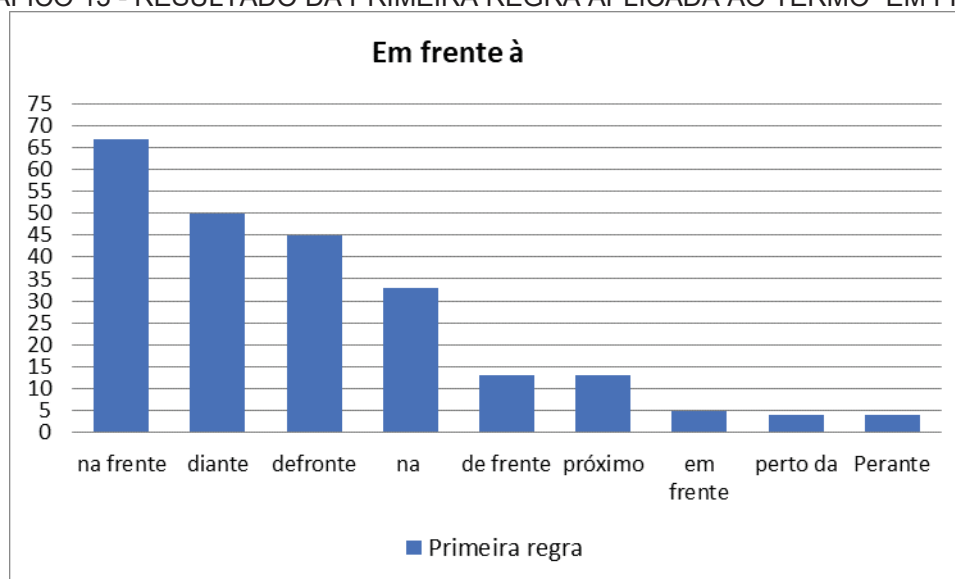
4.5 APLICAÇÃO DE REGRAS PARA CONTRIBUIR NA CRIAÇÃO DE AGRUPAMENTOS DE RELAÇÕES ESPACIAIS SEMANTICAMENTE SIMILARES

A partir dos agrupamentos preliminares da etapa anterior, teve início a aplicação das cinco regras elaboradas na metodologia, conforme o fluxograma proposto.

4.5.1 Primeira Regra

Para que os termos citados poucas vezes não interferissem nos resultados de cada grupo, foram excluídos os termos citados menos do que 1% do total de respostas de cada agrupamento inicial. Como para o termo “em frente à” foram consideradas 267 respostas, arredondando para maior, os termos citados 3 vezes ou menos foram descartados. O GRÁFICO 13 mostra os resultados encontrados para a relação espacial “em frente à”.

GRÁFICO 13 - RESULTADO DA PRIMEIRA REGRA APLICADA AO TERMO “EM FRENTE”

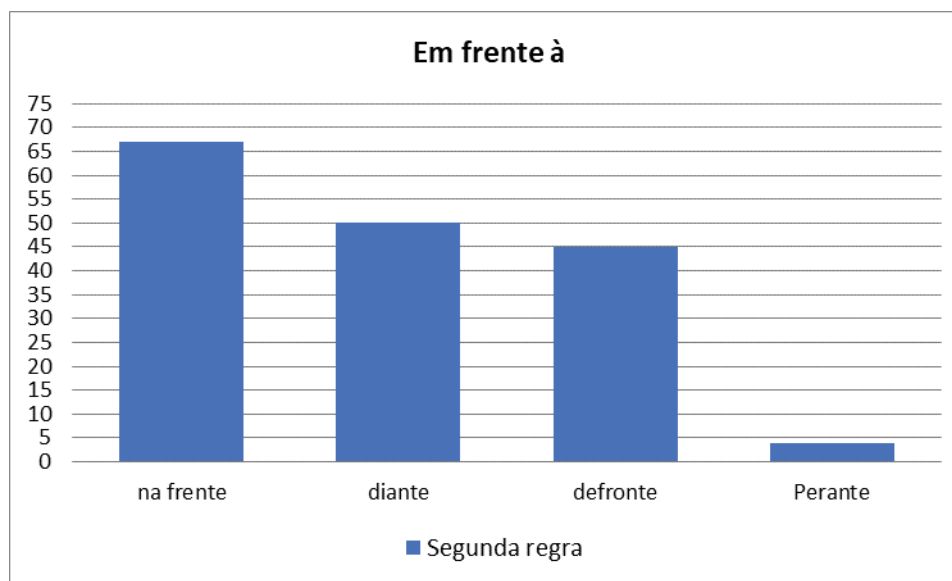


FONTE: A autora (2021).

4.5.2 Segunda Regra

Os termos de relações espaciais que coincidiam com os 14 termos de relações espaciais das 14 frases foram excluídos, pois a versatilidade das relações espaciais ficou evidenciada nas respostas em linguagem natural (FERREIRA; DELAZARI, 2019b) e para este experimento o desejado são os termos que caracterizaram similaridade e não a concorrência entre os 14 termos principais, ou a aparição de termos de relações espaciais vagos que não agregam qualidade às similaridades encontradas. Para o termo “em frente à” foram excluídos os termos “na, de frente a/à/com/para, próximo, em frente à/da/a, e perto da” porque estão nas sentenças principais 2, 14, 6, 1 e 8 com seus próprios resultados. O GRÁFICO 14 mostra os resultados para o termo “em frente à” depois da aplicação desta regra.

GRÁFICO 14 - RESULTADO DA SEGUNDA REGRA APLICADA AO GRUPO “EM FRENTE À”



FONTE: A autora (2021).

Ao final das duas primeiras regras foi possível obter os termos que já se apresentam como similares individualmente por grupo, conforme ilustra a TABELA 18.

TABELA 18 – RESULTADO DAS DUAS PRIMEIRAS REGRAS (continua)

Termos principais	Termos encontrados	Total
em frente à	na frente, na frente da diante, diante da/a/à de frente, de frente a/para, defronte, defronte a/à/da Perante	67 50 45 4
de frente ao	diante, diante ao, diante do na frente, na frente do de frente, de frente ao, defronte, defronte ao, defronte do olhando, olhando o, olhando para perante, perate o de cara, de cara com o, com a cara no	52 45 29 10 5 4
à frente de	adiante, à diante, a diante na frente ali de frente, defronte logo ali aqui diante, diante de, diante do a seguir	69 23 18 13 11 10 10 5
na	numa a, á, à junto a, junto à, junto da no meio, no meio da sobre, sobre a aqui na	35 9 4 4 4 4
em um	num dentro, dentro de/do	126 18

TABELA 18 – RESULTADO DAS DUAS PRIMEIRAS REGRAS

(continua)

Termos principais	Termos encontrados	Total
próximo	na proximidade, nas proximidades, nas proximidades de	16
	a cerca de, acerca, acerca de, cerca de, cerca do	7
	aos arredores de, nos arredores de	5
	nas redondezas do, nas redondezas de, nas redondezas	5
	chegando, chegando a, chegando em	4
perto	na proximidade do, nas proximidades do	6
que tem	onde há	56
	que possuem, que possui	22
	onde existem	17
	que há	15
	de	10
	em que há	9
	na qual existe, na qual existem, a qual existem	7
	na qual há, a qual há	6
	onde tem	6
	que contem, que contém	6
	onde, onde a	5
contendo	4	
com	de	143
	que possui	8
	junto, junto a, junto de	7
	onde há	5
	que contem	4
há	existe, existem	157
ao lado de	na lateral, na lateral de, na lateral em, lateral, à lateral de	34
	junto, junto a, junto ao, junto de	16
	a direita, a direita de, à direita de, a esquerda, à esquerda do	9
	ao redor, ao redor de	8
	paralela ao, paralelo, paralelo a	7
	adjacente, adjacente a	5
	encostada, encostada a, encostado, encostado a	4
	rente, rente a, rente à, renta há	4
	vizinho a, vizinho à, vizinho de	4
	do outro lado de	atravessando, atravessando a
do/ao/no lado oposto, do/no lado oposto da/à, lado oposto da		27
no outro lado, no outro lado da, noutro lado		23
do lado de lá, do lado de lá da		8
nessa, nesta		7
cruzando a, cruzando-se a		6
ali		5
do lado contrário, do lado contrário da, no lado contrário da, ao lado contrário		5
passando a		5
a direita da, a esquerda da, a sua esquerda na		5
daquele lado, daquele lado da		4
naquele lado, naquele lado da	4	
atrás de	às costas, às minhas costas, em minhas costas, nas minhas costas	74
	a/em/na minha retaguarda, na retaguarda	24
	aqui atrás	15
	de trás, detrás, detrás de mim	12
	posterior, posterior à, posterior à mim, posteriormente à mim	12
	para trás, para trás de mim, prá trás	8
	ali atrás	8
	logo atrás, logo atrás de mim	6

TABELA 18 – RESULTADO DAS DUAS PRIMEIRAS REGRAS (conclusão)

Termos principais	Termos encontrados	Total
atrás de	há minha traseira, na minha traseira	6
	ao fundo, aos fundos	5
	por trás, por trás de mim	5
	depois, depois de mim	4
	na parte posterior, na parte posterior a mim	4
entre	no meio, no meio da, no meio de, meio a	94
	em meio, em meio a, em meio à	22
	cercado pela, cercado por	7
	dentre	7
	junto, junto à, junto da	7
	equidistante, equidistante à	4
	no meio entre	4
	a meio caminho, a meio caminho da	4

FONTE: A autora (2021).

4.5.3 Terceira Regra

Esta regra impôs que os pares de grupos que obtiveram respostas com algum termo de relação espacial semelhante fossem unidos. Portanto, sua aplicação resultou na união de nove grupos que apresentaram respostas com termos semelhantes, são eles: “em frente à e de frente ao”, “com e que tem”, “em frente à e à frente de”, “de frente ao e à frente de”, “no, na e em um, em uma”, “do outro lado e à frente de”, “perto e próximo”, “no/na e entre”, “do outro lado da e ao lado de”. Portanto, estes nove grupos, foram unidos e passaram para a quarta e quinta regras, conforme o fluxograma das regras apresentado na metodologia.

Os outros dois grupos de termos não apresentaram respostas iguais e permaneceram como estavam após a aplicação da segunda regra: “há e atrás”, portanto, eles passaram diretamente para a quinta regra.

Iniciando a aplicação da terceira regra pelo termo “em frente à” percebeu-se que este termo apresentou alguns dos mesmos termos mencionados na tabela de termos “de frente ao”, são eles: “diante, na frente, de frente, perante” (ilustrados em azul na tabela). Portanto estes dois grupos foram unidos (TABELA 19), conforme determinou a terceira regra. Todas as respostas foram consideradas e agrupadas mesmo estando com a grafia incorreta por se tratar de uma pesquisa em linguagem natural.

TABELA 19 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “EM FRENTE À E DE FRENTE AO”

Termo	Termos encontrados	Citações
em frente à	na frente, na frente da	67
	diante, diante da/a/à	50
	defronte, defronte a/à/da/para	45
	perante	4
de frente ao	na frente, na frente do	45
	diante, diante ao/do	52
	defronte, defronte ao/do	29
	perante, perante o	5
	olhando, olhando o/para	10
de cara, de cara com, com a cara	4	

FONTE: A autora (2021).

Continuando com o termo “em frente à” percebeu-se a existência de mais um grupo de termos com alguns termos de respostas semelhantes, o grupo “à frente de”. Os termos coincidentes foram: “na frente, diante, de frente”, portanto, estes dois grupos também foram unidos, conforme TABELA 20:

TABELA 20 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “EM FRENTE À E À FRENTE DE”

Termo	Termos encontrados	Citações
em frente à	na frente, na frente da	67
	diante, diante da/a/à	50
	defronte, defronte a/à/da/para	45
	perante	4
à frente de	adiante, à diante, a diante	69
	na frente	23
	ai	18
	de frente, defronte	13
	logo ali	11
	aqui	10
	diante, diante de, diante do	10
	a seguir	5

FONTE: A autora (2021).

Passando a análise para o termo “na”, segundo termo pela ordem das sentenças, percebeu-se que existiram respostas semelhantes nos resultados da segunda regra para o termo “em um” (TABELA 21) e para o termo “entre” (TABELA 22). Entendem-se aqui pelas regras da língua portuguesa brasileira que para o termo “na” também deverá ser considerado o termo “no” e para o termo “em um” também deverá ser considerado o seu termo correspondente feminino. A aplicação do gênero para os termos de relações espaciais foi levada em consideração porque a gramática da língua portuguesa brasileira exige isto. Nestes dois grupos foram

considerados como semelhantes os termos “numa” e “num”, o que possibilitou a união destes grupos.

TABELA 21 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “NA E EM UM”

Termo	Termos encontrados	Citações
na	numa	35
	a, á, à	9
	junto a, junto à, junto da	4
	no meio, no meio da	4
	sobre, sobre a aqui na	4
em um	num dentro, dentro de/do	126 18

FONTE: A autora (2021).

TABELA 22 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “NA E ENTRE”

Termo	Termos encontrados	Citações
Na	numa	35
	a, á, à	9
	junto a, junto à, junto da	4
	no meio, no meio da	4
	sobre, sobre a aqui na	4
entre	no meio, no meio da, no meio de	90
	em meio, em meio a, em meio à	22
	cercado pela, cercado por	7
	dentre	7
	junto, junto à, junto da	7
	equidistante, equidistante à	4
	no meio entre	4
	no meio, no meio da, meio a	4
a meio caminho, a meio caminho da	4	

FONTE: A autora (2021).

Segue a TABELA 23 com o termo “próximo” que foi unido ao grupo de termos “perto”.

TABELA 23 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “PERTO DO” E “PRÓXIMO DE”

Termo	Termos encontrados	Citações
próximo de	na proximidade, nas proximidades, nas proximidades de	16
	a cerca de, acerca, acerca de, cerca de, cerca do	7
	aos arredores de, nos arredores de	5
	nas redondezas do, nas redondezas de, nas redondezas	5
	chegando, chegando a, chegando em	4
perto de	na proximidade do, nas proximidades do	6

FONTE: A autora (2021).

Os grupos “com” e “que tem” (TABELA 24), “de frente ao e à frente de” (TABELA 25), “do outro lado da e à frente de” (TABELA 26) e “do outro lado e ao lado de” (TABELA 27) apresentaram semelhanças nas respostas e foram unidos.

TABELA 24 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “COM” E “QUE TEM”

Termo	Termos encontrados	Citações
com	de	143
	que possui	8
	junto, junto a, junto de	7
	onde há	5
	que contem	4
que tem	onde há	56
	que possuem, que possui	22
	onde existem	17
	que há	15
	de	10
	em que há	9
	na qual existe, na qual existem, a qual existem	7
	na qual há, a qual há	6
	onde tem	6
	que contem, que contém	6
	onde, onde a	5
	contendo	4

FONTE: A autora (2021).

TABELA 25 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “DE FRENTE AO E À FRENTE DE”

Termo	Termos encontrados	Citações
de frente ao	na frente, na frente do	45
	diante, diante ao, diante do	52
	de frente, de frente ao, defronte, defronte ao, defronte do	29
	perante, perante o	5
	olhando, olhando diretamente para, olhando o, olhando para, olhando para o	10
	de cara, de cara com o, com a cara no	4
à frente de	Adiante, À diante, a diante	69
	Na frente	23
	Ali	18
	de frente, defronte	13
	Logo ali	11
	Aqui	10
	diante, diante de, diante do	10
	A seguir	5

FONTE: A autora (2021).

TABELA 26 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “DO OUTRO LADO E À FRENTE DE”

Termo	Termos encontrados	Citações
do outro lado	atravessando, atravessando a	58
	do lado oposto, do lado oposto da, lado oposto da, ao lado oposto, no lado oposto, no lado oposto à, no lado oposto da	27
	no outro lado, no outro lado da, noutro lado	23
	do lado de lá, do lado de lá da	8
	nessa, nesta	7
	cruzando a, cruzando-se a	6
	ali	5
	do lado contrário, do lado contrário da, no lado contrário da, ao lado contrário	5
	passando a	5
	a direita da, a esquerda da, a sua esquerda na	5
	daquele lado, daquele lado da	4
naquele lado, naquele lado da	4	
à frente de	adiante, à diante, a diante	69
	na frente	23
	ali	18
	de frente, defronte	13
	logo ali	11
	aqui	10
	diante, diante de, diante do	10
a seguir	5	

FONTE: A autora (2021).

TABELA 27 - RESULTADO DA TERCEIRA REGRA APLICADA AOS GRUPOS “DO OUTRO LADO E AO LADO DE”

Termo	Termos encontrados	Citações
do outro lado	atravessando, atravessando a	58
	do lado oposto, do lado oposto da, lado oposto da, ao lado oposto, no lado oposto, no lado oposto à, no lado oposto da	27
	no outro lado, no outro lado da, noutro lado	23
	do lado de lá, do lado de lá da	8
	nessa, nesta	7
	cruzando a, cruzando-se a	6
	ali	5
	do lado contrário, do lado contrário da, no lado contrário da, ao lado contrário	5
	passando a	5
	a direita da, a esquerda da, a sua esquerda na	5
	daquele lado, daquele lado da	4
naquele lado, naquele lado da	4	
ao lado de	na lateral, na lateral de, na lateral em, lateral, à lateral de	34
	junto, junto a, junto ao, junto de	16
	a direita, a direita de, à direita de, a esquerda, à esquerda do	9
	ao redor, ao redor de	8
	paralela ao, paralelo, paralelo a	7
	adjacente, adjacente a	5
	encostada, encostada a, encostado, encostado a	4
	rente, rente a, rente à, renta há	4
	vizinho a, vizinho à, vizinho de	4

FONTE: A autora (2021).

4.5.4 Quarta Regra

Nesta regra foram considerados apenas os termos citados nos dois grupos unidos (ilustrados em azul) a partir da terceira regra tendo suas citações somadas, e os termos que foram citados em apenas um dos grupos foram eliminados. No grupo "de frente ao" foram citados os termos "olhando, olhando o, olhando para" dez vezes e os termos "De cara, de cara com, com a cara" foram citados quatro vezes, mas esses termos não foram citados no grupo "em frente à", portanto esses termos foram excluídos, conforme TABELA 28.

TABELA 28 - RESULTADO DA QUARTA REGRA APLICADA AOS GRUPOS "EM FRENTE À E DE FRENTE AO"

Termos	Termos encontrados	Citações
em frente à + de frente ao	na frente, na frente da	112
	diante, diante da/a/à	102
	defronte, defronte	74
	à/ao/do/da/para	9
	perante, perante o	eliminado
	olhando, olhando o/para	eliminado
	de cara, de cara com, com a cara	eliminado

FONTE: A autora (2021).

Para os termos "em frente à e de frente ao" a TABELA 29 ilustra os resultados da quarta regra.

TABELA 29 - RESULTADO DA QUARTA REGRA APLICADA AOS GRUPOS "EM FRENTE À E À FRENTE DE"

Termos	Termos encontrados	Citações
em frente à + à frente de	na frente, na frente da	90
	diante, diante da/a/à	60
	de frente, de frente a/para, defronte, defronte	58
	a/à/da	eliminado
	perante	eliminado
	adiante, à diante, a diante	eliminado
	ali	eliminado
	logo ali	eliminado
	aqui	eliminado
	a seguir	eliminado

FONTE: A autora (2021).

Seguindo o raciocínio da quarta regra a TABELA 30 mostra o resultado para todos os grupos que foram unidos na terceira regra.

TABELA 30 - RESULTADO DA QUARTA REGRA APLICADA À UNIÃO DOS GRUPOS

Termos principais	Termos encontrados	Total
em frente à + de frente ao	na frente, na frente da	112
	diante, diante da/a/à	102
	defronte, defronte à/ao/do/da/para	74
	perante, perante o	9
em frente à + à frente de	na frente, na frente da	90
	diante, diante da/a/à	60
	de frente, de frente a/para, defronte, defronte a/à/da	58
na + em um	numa	35
	num	126
na + entre	junto a, junto à, junto da	11
	no meio, no meio da	8
próximo de + perto de	na proximidade, nas proximidades, nas proximidades de/do	22
com + que tem	de	153
	que possuem, que possui	30
	onde há	61
	que contem, que contém	10
de frente ao + à frente de	na frente, na frente do	68
	diante, diante ao, diante do	62
	de frente, de frente ao, defronte, defronte ao, defronte do	42
do outro lado + à frente de	ali	23
do outro lado + ao lado de	a direita da, a esquerda da, a sua esquerda na	14

FONTE: A autora (2021).

4.5.5 Quinta Regra

A partir dos resultados da quarta regra, os termos que não apresentam citações maiores que 10% do termo mais citado de cada grupo devem ser excluídos. Como o objetivo do trabalho é trazer os termos semanticamente similares para que sejam utilizados em PLN e/ou em aplicações de geoprocessamento, o importante é que fiquem os termos que realmente são mais utilizados e são entendíveis por todas as pessoas. Os resultados da quinta regra, para todos os grupos, estão ilustrados na TABELA 31. Nesta tabela também consta a classificação dos termos encontrados na aplicação das regras.

TABELA 31 - RESULTADO DA QUINTA REGRA APLICADA À UNIÃO DOS GRUPOS

Termos principais	Termos semanticamente similares encontrados pelas regras	Classificação	Citações	Total
em frente à, de frente ao	na frente, na frente da/do/de diante, diante da/do/de/a/à/ao defronte, defronte à/ao/do/da/de/para	projetiva projetiva projetiva	112 102 74	288
com, que tem	de que possuem, que possui onde há	preposição com sentido de posse topológica verbo	153 30 61	244
em frente à, à frente de	na frente, na frente da/do/de diante, diante da/a/à de frente, de frente a/para, defronte, defronte a/à/da	projetiva projetiva projetiva	90 60 58	208
de frente ao, à frente de	na frente, na frente da/do/de diante, diante da/do/de/a/à/ao de frente, de frente ao, defronte, defronte ao, defronte do	projetiva projetiva projetiva	68 62 42	172
no/na, em um/uma	num/numa	topológico	161	161
há	existe, existem	verbo	157	157
atrás	às/em/nas minhas costas, às costas a/em/na minha retaguarda, na retaguarda aqui atrás de trás, detrás, detrás de mim posterior, posterior à/ à mim, posteriormente à mim para trás, para trás de mim, prá trás ali atrás	projetiva projetiva dêitica/projetiva projetiva projetiva projetiva dêitica/projetiva	74 24 15 12 12 8 8	153
do outro lado da, à frente de	ali	dêitica	23	23
próximo de, perto de	na proximidade do, nas proximidades do	distância	22	22
no/na, entre	junto a, junto à, junto da no meio, no meio da	adjacente topológico	11 8	19
do outro lado da, ao lado de	à direita/esquerda, à direita/esquerda de/da/do	projetiva	14	14

FONTE: A autora (2021).

Sobre os resultados encontrados após a aplicação das cinco regras deve ser realizada uma análise (por uma pessoa nativa da língua portuguesa) dos grupos de termos para ver a coerência dos termos e a verificação da gramática da língua em questão.

A língua portuguesa brasileira faz distinção do gênero feminino e masculino, do número singular e plural, artigos definidos e indefinidos, e das contrações e combinações de preposições. Por isto, ao analisar os termos “em frente”, “de frente”, “na frente” e “à frente”, percebe-se que o vocábulo “frente” não muda, porém, os

termos que caracterizam estas relações espaciais podem ou não ser diferentes. O exemplo que melhor justifica é o da relação espacial “à frente de” que foi utilizado na frase 15 (À FRENTE de onde estou situado está o Centro Histórico da cidade.) que se for analisado até a segunda regra trás termos relacionados à: a) na dianteira; b) no comando (“Michaelis Moderno Dicionário Inglês & Português”, 2021); e a) na parte dianteira; contrário à traseira: o carro estava à frente da igreja; b) posição vantajosa em relação aos demais: o time está à frente no campeonato c) avante; seguido o comando, à direção (FERREIRA, 2010). Porém se for analisado até a quinta regra traz termos semanticamente similares aos termos “em frente à e de frente à”.

Os resultados mostraram que para o termo “à frente de” dos 159 termos filtrados pela segunda regra, os termos que representaram o que os dicionários apresentam foram: “adiante” com 69 citações, “ali” com 18 citações, “logo ali” com 11 citações e “a seguir” com 5 citações, ou seja, 65% das citações. No dia a dia em linguagem natural as pessoas não conseguiram utilizar o termo “à frente de” 100% como deveria ser. E os 35% mencionados (na frente, defronte e diante) como similares aos termos “em frente à e de frente ao” fizeram com que as regras criadas o classificassem como semanticamente similares a estes termos. Aqui entra a importância da análise dos resultados ser executada por um nativo da língua em questão, que pode ou não acatar os resultados das regras elaboradas.

Pode-se considerar como termos semanticamente similares as relações espaciais “em frente à” e “de frente ao” que com os complementos destes vocábulos representam a mesma posição do *locatum* em relação ao *relatum*. E os termos e “na frente”, “diante” e “defronte” também podem ser considerados semanticamente similares porque além de terem sido citados nos dois grupos, não alteraram o significado e nem a posição geográfica da situação envolvida. Os termos semanticamente similares encontrados para estes dois termos não devem ser considerados similares para o termo “à frente” que significa que o lugar da sentença (Centro Histórico) está adiante do lugar onde a pessoa se encontra localizada.

Outro fator importante é que foi utilizado o egocentrismo como eixo de referência neste teste, portanto, assume-se que “em frente” e “de frente” devem ser considerados similares porque apresentaram termos semanticamente similares e espacialmente significam que estão “na frente” do lugar propriamente dito. Isto pôde ser verificado pelo termo “perante” que foi utilizado nestes dois grupos, mas que foi

excluído quando a quinta regra foi aplicada, esta regra que determina que sejam considerados somente os termos mais relevantes (pelo menos 10% de citações do primeiro colocado do grupo).

Os três termos “em frente, de frente e à frente” trazem uma confusão quando utilizados em linguagem natural, como se percebe pelos resultados encontrados na TABELA 30, onde os voluntários citaram os termos “na frente”, “de frente” e “diante” como similares para os três termos. Porém estes termos serão considerados similares para “em frente” e “de frente”. Isto ocorre pelo fato de o termo “à frente” ser considerado como contrário do termo “atrás”, quando considerados termos que possuem “frente” e “trás”, como, por exemplo: casas, carros, entre outros. Os termos semanticamente similares encontrados provam que na linguagem natural do cotidiano das pessoas, situações de localização e posição espacial geralmente não levam em conta somente as regras da língua portuguesa brasileira e sim as experiências e as percepções dos ambientes vividas por cada indivíduo, concordando com (STOCK et al., 2010).

Os termos “com” e “que tem” apresentaram resultados similares e são termos que mostram detalhes do lugar ou do *relatum* e desta maneira contribuem com a localização/posição do *locatum* por detalhamento. Os termos deste grupo aparentemente não são considerados como relações espaciais, porém, eles vêm sendo utilizados como tal com muita frequência em descrições em linguagem natural de localização/posição. Os termos “de”, “onde há” e “que possui, que possuem” foram encontrados como semânticos destes dois grupos como mostra a TABELA 30.

Os termos “no, na, em um, em uma” mostraram-se similares entre si não havendo distinção na contração dos artigos definidos e indefinidos, e trouxeram como semanticamente similares os termos “num” e “numa” (TABELA 30), que são as contrações dos termos “em um” e “em uma”. Isto mostra que em descrições de posicionamento/localização em linguagem natural do dia a dia, as pessoas não fazem a distinção de artigo definido e indefinido para os elementos envolvidos na situação geográfica (língua do Português Brasileiro).

O termo no/na também foi unido ao termo “entre” pelos termos “junto e no meio” que são das classes adjacente e topológica, respectivamente. Esta junção não foi muito significativa porque teve apenas 19 citações, sendo 11 para o termo “junto” e 8 para o termo “no meio”. E o termo entre geralmente deve ser acompanhado de dois *relata* ou de um *relatum* no plural, diferentemente do termo que seria

desmembrado em duas expressões locativas se tivesse como *relata* dois elementos diferentes.

O termo “do outro lado da” foi unido ao termo “à frente de” pelo termo “ali” e ao termo “ao lado de” pelos termos “à direita/esquerda”. Porém o termo “ali” no grupo “do outro lado da” foi mencionado apenas 5 vezes, ou seja, menos de 10% do primeiro termo mais mencionado na segunda regra que foi o verbo “atravessando” com 58 citações. E o mesmo aconteceu com os termos “à direita/esquerda”. Em linguagem natural estes casos ocorreram, mas não são de extrema importância como similares do termo “do outro lado” apesar de serem da classe projetiva.

Os termos “perto” e “próximo” são classificados pela literatura como sendo termos de distância e foram citados pelos voluntários nos agrupamentos iniciais como respostas nas frases dos termos: “em frente”, “de frente”, “entre”, “atrás”, “à frente”, “ao lado”, “no/na”, “em um/uma”. Isto mostra que são termos muito utilizados, porém são termos vagos e que não agregam informação precisa de localização/posição por não se ter uma definição de quão perto se está dos elementos (DERUNGS; PURVES, 2016; STOCK et al., 2010). Os termos “perto” e “próximo” foram utilizados em quase todos os grupos e não contribuem ou contribuem pouco para programas de localização do tipo *Where am I?*. De acordo com (HALL; JONES, 2008) muitos aspectos da linguagem espacial relacionados às relações entre entidades espaciais são vagos e as tecnologias GIS fornecem pouco suporte para lidar com essa imprecisão, em parte, por falta de modelos quantitativos para relações espaciais vagas. Porém em PNL podem ser utilizados como semanticamente similares.

Os termos que não apresentaram semelhanças nas respostas e não passaram pela terceira e quarta regras, também passaram pela quinta regra para que fossem filtrados pelos seus termos mais relevantes: “há” e “atrás”.

O termo “há” tem como similaridade os termos “existe” e “tem”, conforme a gramática da língua portuguesa brasileira. E são termos que estão sendo cada vez mais utilizados como relações espaciais em descrições de localização, sendo baseados na existência de pontos de referência na região da descrição.

Os sujeitos mencionaram os termos “direita” e “esquerda” pensando em “lado direito” e/ou “lado esquerdo”, apesar do ângulo de visão do observador e do eixo de referência não terem sido fixados. Portanto, nas análises subentendeu-se que estes termos representavam lateralidade referente ao *locatum*. São termos

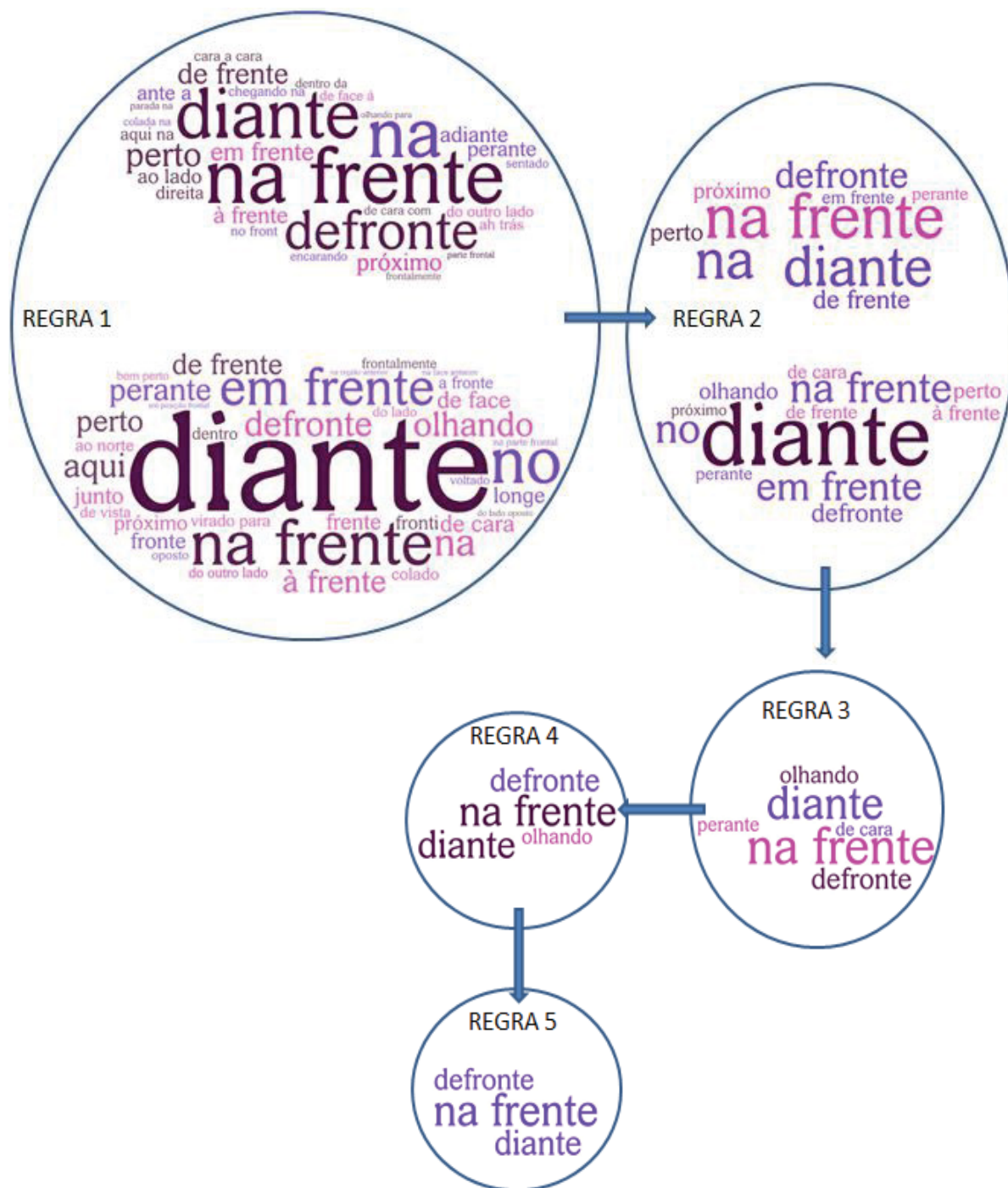
comumente utilizados em descrições de localização em linguagem natural, por isto eles foram deixados como similares de “ao lado de”.

Os termos “posterior, posterior à, posterior à mim, posteriormente à mim” encontrados como similares de “atrás de mim” neste teste, se contrapõem ao estudo de Teixeira (2013) que trouxe como resultado o termo “anterior à”. Para esta divergência podem existir dois fatores que precisam de mais análises, o primeiro é que os estudos de Teixeira (2013) são na língua portuguesa europeia e o segundo é que ele tenta distinguir termos utilizados como equivalentes entre tempo e espaço, ou seja, ele afirma que espacialmente o termo “atrás” está relacionado ao tempo passado e corresponde aos termos temporais de “antes” e “anterioridade”; e que o termo espacial “à frente” está relacionado ao tempo futuro e corresponde aos termos “depois” e “posterioridade”. O autor ainda afirma que para os falantes, as equivalências julgadas pertinentes na linguagem natural são exatamente opostas às que os dicionários prioritariamente apresentam, porque os dicionários identificam um termo espacial como um termo temporal. Porém nesta tese, não estão sendo diferenciados termos de relações espaciais no aspecto temporal e sim, no aspecto geral de utilização em linguagem natural.

Nos resultados dos agrupamentos após a aplicação das cinco regras percebeu-se que a quantidade de termos foi diminuindo a cada regra aplicada até que restassem os termos considerados como semanticamente similares, conforme mostra o exemplo da FIGURA 27 para os termos “em frente” e “de frente” que passaram por todas as regras. Os termos citados para cada grupo partiram dos termos principais, no caso do exemplo para os termos “em frente” e “de frente”, depois de passarem pelas duas primeiras regras, houve a união (terceira regra) e a seleção dos termos comuns aos dois grupos (quarta regra) e a seleção dos termos mais relevantes (quinta regra).

A aplicação das regras foi um importante passo para se chegar aos grupos semanticamente similares contrastando com o trabalho de Marcińczuk, Oleksy e Wieczorek (2016) que ainda tem que resolver o problema da diversidade de termos de relações espaciais encontrados no reconhecimento de expressões espaciais na língua polonesa.

FIGURA 27 - RESULTADO DA APLICAÇÃO DE TODAS AS REGRAS PARA OS TERMOS "EM FRENTE" E "DE FRENTE"



FONTE: A autora (2021).

4.5.6 Agrupamentos finais

Depois da aplicação das regras, foi possível obter os agrupamentos finais de termos semanticamente similares (TABELA 32), corroborando com Ferreira e Delazari (2019a) que explicam que embora relações espaciais usadas erroneamente

causem entendimentos semelhantes às pessoas, isso não implica que a escolha de qualquer uma delas denote a mesma informação espacial. Os agrupamentos da TABELA 32 foram definidos depois da análise principalmente das junções da terceira regra.

TABELA 32 - TERMOS SEMANTICAMENTE SIMILARES ANALISADOS POR NATIVO DA LÍNGUA

Termos principais	Termos semanticamente similares encontrados por classes	Classificação	Citações	Total
em frente à, de frente ao	na frente, na frente da/do/de	projetiva	112	288
	diante, diante da/do/de/a/à/ao	projetiva	102	
	defronte, defronte à/ao/do/da/de/para	projetiva	74	
no/na, em um/uma	num/numa	topológico	161	161
há	existe, existem	verbo	157	157
entre	no meio, no meio da, no meio de, meio a	topológico	94	134
	em meio, em meio a, em meio à	topológico	22	
	cercado pela, cercado por	topológico	7	
	dentre	topológico	7	
atrás	no meio entre	topológico	4	130
	às/em/nas minhas costas, às costas	projetiva	74	
	a/em/na minha retaguarda, na retaguarda	projetiva	24	
	de trás, detrás, detrás de mim	projetiva	12	
	posterior, posterior à/ à mim, posteriormente à mim	projetiva	12	
para trás, para trás de mim, prá trás	projetiva	8		
à frente de	adiante, à diante, a diante	projetiva	69	69
do outro lado de	do/ao/no lado oposto, do/no lado oposto da/à, lado oposto da	projetiva	27	60
	no outro lado, no outro lado da, noutro lado	projetiva	23	
	do lado contrário, do lado contrário da, no lado contrário da, ao lado contrário	projetiva	5	
ao lado de	a direita da, a esquerda da, a sua esquerda na lateral, na lateral de, na lateral em, lateral, à lateral de	projetiva	34	43
	a direita, a direita de, à direita de, a esquerda, à esquerda do	projetiva	9	
com, que tem	que possuem, que possui	topológica	30	30
próximo de, perto de	na proximidade do, nas proximidades do	distância	22	22

FONTE: A autora (2021).

Os advérbios de lugar “lá, ali e aqui” e os pronomes demonstrativos “essa” e “esta” que aparecem através da contração da preposição “em” formando os termos “nessa e nesta” (TABELA 31) foram excluídos dos resultados. Isto porque são considerados termos de distância onde os elementos podem estar “perto” ou “longe” do falante e porque são termos vagos. Eles são úteis quando utilizados em diálogos presenciais onde os sujeitos falantes estão *in loco* e podem apontar para os elementos utilizados como referência (JUNGBLUTH, 2005; PEREIRA, 2009; TEIXEIRA, 2005). Por isto não contribuem espacialmente com a posição/localização

dos *relata* em relação ao *locatum* neste experimento que trás termos que podem ser aplicados computacionalmente.

Os verbos “atravessando”, “cruzando”, “passando”, “a seguir” e “encostado” apesar de terem sido mencionados como semanticamente similares para alguns dos termos principais, foram excluídos por serem verbos de movimento e estarem fora do contexto principal desta tese.

E para que os resultados tivessem as classificações correspondentes às classes dos termos principais, todos os termos de classificações diferentes foram excluídos. Os termos “com e que tem” que são os termos principais permaneceram porque são classificados como topológicos.

Além das diferenças nas classes utilizadas por diferentes autores, existem diferenças nas classes às quais os termos de relação espacial individual são atribuídos. Portanto, a classificação dos termos de relação espacial foi realizada de acordo com os esquemas abaixo:

- **Projetiva:** esta classe está incluída em quase todos os esquemas anteriores, embora às vezes subdividida em horizontal e vertical (KEMMERER, 2006), são combinadas com classes direcionais ou dimensionais (RETZ-SCHMIDT, 1988; TENBRINK, 2007).
- **Topológica:** também aparece em quase todos os outros esquemas. Consideramos a classe de contenção de Bitters (2009) uma subclasse de topologia, seguindo definições comuns de relações topológicas (COHN et al., 1997).
- **Proximidade:** é identificada como uma classe separada em uma série de esquemas (BITTERS, 2009; COVENTRY; GARROD, 2004), como uma classe dêitica (KEMMERER, 2006), e a classe relacionada à distância de Tenbrink (2007) que incluem relações de proximidade.
- A classe Direcional de Zwarts (2005) e Zwarts e Winter (2000) contêm termos de relação espacial semelhantes à classe relacionada ao percurso de Tenbrink (2007) que foi adotada com o termo Direcional.
- Também foram adotadas três classes identificadas em trabalhos anteriores que apresentam menos consenso, mas que refletem relações que não estão listadas como incluídas nas classes acima, sendo estas a

Adjacência (BITTERS, 2009); Direções Cardeais (COVENTRY; GARROD, 2004) e Orientação (BITTERS, 2009).

Os resultados mostraram que as relações espaciais e seus termos semanticamente similares quando agrupados e contados assumem uma ordem de preferência na sua utilização conforme mostra a TABELA 32. E mais uma vez a classe projetiva através dos termos “em frente à/ao/da/do/de” e “de frente à/ao/da/do/de” com seus termos similares encontrados contam com o primeiro lugar em preferência para a língua portuguesa brasileira.

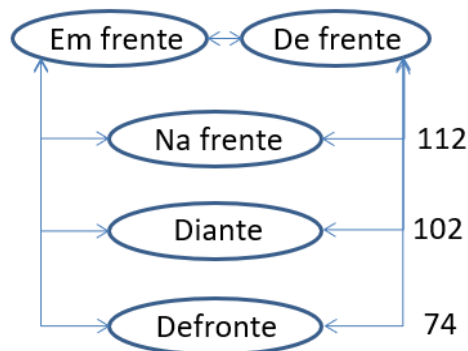
De acordo com Fagundes et al. (2021) os resultados do maior uso de termos da classe projetiva para a língua portuguesa brasileira podem ser associados a descrições de visão aérea, como se os falantes estivessem analisando um mapa visto de cima (*survey perspective*) (NOORDZIJ; ZUIDHOEK; POSTMA, 2006). Em contraste com a língua inglesa que utiliza com mais frequência termos da classe direcional, como se o observador estivesse se movendo traçando uma rota (*route perspective*) (DENIS, 1997; NOORDZIJ; ZUIDHOEK; POSTMA, 2006), e muitas vezes empregando o movimento fictício (TALMY, 2000).

Quando as classificações dos grupos que foram encontrados como semanticamente similares foram analisadas percebeu-se que a classe projetiva se sobressaiu sobre as demais classes. De acordo com Hall e Jones (2008) e Pinker e Ravagnani (2008), ainda há muito a ser estudado, analisado e compreendido ao lidar com experimentos com descrições de seres humanos em linguagem natural porque o que vem sendo realizado e analisado são modelos com base em dados da *WEB* que possibilita a coleta de um número maior de amostras. Experimentos com seres humanos limitam as amostras em um número muito menor que a *WEB*. E quando as pessoas usam a linguagem natural diária para expressar sua localização, há uma tendência para o uso mais frequente de certos termos de relações espaciais, apesar da liberdade de expressão que a fala e a escrita proporcionam, como é o caso da classe direcional na língua inglesa (FAGUNDES et al., NO PRELO). Além disso, a linguagem natural permite que os usuários façam combinações de palavras e/ou usem palavras semelhantes para explicar sua localização. As relações espaciais em qualquer linguagem natural têm um número limitado de palavras, mas a linguagem natural permite que as pessoas usem termos que não são espaciais em situações espaciais, tornando-as também relações espaciais.

4.5.7 Distância entre os termos semanticamente similares

De acordo com os resultados apresentados, os termos encontrados foram agrupados conforme a quantidade de vezes que foram citados pelos voluntários. Desta maneira, atribuiu-se um distanciamento aos termos de cada grupo encontrado pela ordem decrescente das citações, sendo que, o termo mais próximo ao termo principal, ou seja, o termo com mais citações é o termo similar mais utilizado e o termo mais distante ou o menos citado é o termo similar menos utilizado - diferenciando dos estudos de Bruns e Egenhofer (1996) e Cuayáhuitl et al. (2010). Os termos encontrados podem ser substituídos uns pelos outros que serão entendidos e representarão a mesma localização geográfica, porém o distanciamento mostra quais são os termos mais significativos pela ordem decrescente das citações. No caso da FIGURA 28, foram encontrados como similares os termos principais “em frente” e “de frente” e os termos “na frente” com 112 citações, “diante” com 102 citações e “defronte” com 74 citações; o que indica pela quantidade de citações que as similaridades entre eles são dadas pela ordem decrescente. Isto contribui com aplicações computacionais que podem fazer a utilização tanto dos termos mais citados como de todos os termos, de acordo com as similaridades encontradas. Corroborando com March Castañeda (2017) que cita que para a criação de uma ontologia de relações espaciais, o ideal é que sejam utilizados todos os termos semanticamente similares encontrados ou os mais próximos, assim como para o PLN (HU, 2018).

FIGURA 28 - DISTÂNCIA ENTRE OS TERMOS SIMILARES DO AGRUPAMENTO “EM FRENTE” E “DE FRENTE”



FONTE: AUTORA

4.6 VERIFICAÇÃO DAS HIPÓTESES PROPOSTAS PARA O TESTE DE ASSOCIAÇÃO LIVRE

As hipóteses analisadas para este teste foram:

1ª hipótese para o teste de associação livre: Os termos semanticamente similares de relações espaciais encontrados através do mesmo *relatum* e do mesmo *locatum* dos testes “online” (4.1.5) e “presencial” (4.2.5) coincidem com os termos encontrados no teste de “associação livre” (4.5.6)?

Para que esta hipótese fosse testada, foram analisados os termos de relações espaciais encontrados com os mesmos *locata* e *relata* nos 9 lugares do primeiro teste (*outdoor*) e nas duas posições do segundo teste (*indoor*). Todos os termos foram classificados em topológico, projetivo, distância, adjacência, orientação, cardinal e verbos, pois isto facilitou as análises e permitiu a separação dos termos de acordo com a classificação. A TABELA 33 mostra que para o contexto *outdoor* foram encontrados 277 termos de relações espaciais utilizados com os mesmos *locata* e *relata* e que deste total, 50% dos termos coincidiram com os termos encontrados no terceiro teste (associação livre) de acordo com a TABELA 32. No contexto *indoor*, dos 252 termos encontrados, 63% coincidiu com o terceiro teste, conforme TABELA 34.

TABELA 33 - TERMOS SEMANTICAMENTE SIMILARES - *OUTDOOR*

Site	Total	Coincidentes
BRA1	39	23
BRA2	20	9
BRA3	45	30
BRA4	65	29
NZ1	26	17
NZ2	26	4
NZ3	29	15
NZ4	8	4
NZ5	19	7
total	277	138

FONTE: A autora (2021).

TABELA 34 - TERMOS SEMANTICAMENTE SIMILARES - *INDOOR*

Posição	Total	Coincidentes
A	133	87
B	119	72
total	252	159

FONTE: A autora (2021).

Portanto, com o total de termos nos dois contextos e as respectivas coincidências de termos, a hipótese se verifica em 56%. Considerando que os dois primeiros testes não tinham como previsão a análise de termos que fossem similares, o resultado foi promissor, ultrapassando a metade dos termos encontrados com os mesmos *locata* e *relata*. Isto também corrobora com o resultado da hipótese de que na língua portuguesa muitos dos termos são utilizados em diferentes contextos.

Os 44% dos termos que não coincidiram representam que são de classificações diferentes. Destes 44% foram retirados os termos que foram citados e que coincidiam com os 14 termos principais analisados, como foi o caso da segunda regra aplicada no terceiro teste. Os termos restantes são de classificações diferentes, como ilustra a TABELA 35.

TABELA 35 - TERMOS COM CLASSIFICAÇÕES DIFERENTES - *OUTDOOR* E *INDOOR*

Relação espacial	Termos com mesmo <i>relata</i> e <i>locata</i>	Classificação
de/em frente à/ao/da/do/de/para	olhar para, ao redor da onde exatamente estou da para ver consigo ver, vejo subi	direção topológico sem classe distância verbo de movimento
tem	no final oposto, no final uns 10 metros nessa já pode se ver	projetiva distancia topológico distância
próximo	depois da, acima da em direção à	projetiva direção
ao lado	atrás ao norte na direção sul	projetiva cardinal direção
no	aqui que dá para possui anexo à ao longo desta perpendicular na direção da	distância direção verbo adjacência direção orientação direção
com	abaixo do lado de fora	projetiva topológico
entre	divide	verbo de movimento
do outro lado da	contém	verbo

FONTE: Autora (2021).

Existiram descrições em que os voluntários utilizaram três ou quatro termos diferentes de relações espaciais para a mesma situação geográfica, conforme alguns exemplos mostrados nas TABELAS 36 para o contexto *outdoor* e 37 para o contexto *indoor*. Isto acontece com frequência porque não há certo ou errado em se

tratando da escolha do termo de relação espacial, em linguagem natural, e sim o entendimento e a percepção individual de cada um, corroborando com o estudo de (TENBRINK, 2017).

TABELA 36 - TERMOS DIFERENTES DE RELAÇÕES ESPACIAIS PARA OS MESMOS LOCATA E RELATA - OUTDOOR

Localidade	Locatum	Relatum	Relações espaciais
BRA1	eu	praça	(l) estou localizado em uma (r) (l) estou situado perto de (r) (l) estou em frente (r)
BRA2	implícito - acidente	supermercado SUPERPOP	(l implícito - acidente) ao lado do (r) (l implícito - acidente) em frente ao (r) (l implícito - acidente) próximo ao (r)
BRA3	implícito - acidente	praça	(l implícito - acidente) em frente a (r) (l implícito - acidente) próximo a (r) (l implícito - acidente) mais ou menos na metade da extensão da (r)
NZ3	implícito - acidente, <u>acidente</u>	igreja	(l implícito - acidente) próximo do que parece ser (r) (l implícito - acidente) em frente à (r) (l implícito - acidente) perto de (r) (l) em frente da (r) (l implícito - acidente) ao fundo de (r)

FONTE: Autora (2021).

TABELA 37 - TERMOS DIFERENTES DE RELAÇÕES ESPACIAIS PARA OS MESMOS LOCATA E RELATA – INDOOR

Localidade	Locatum	Relatum	Relações espaciais
Posição A	nós, implícito - eu, eu	departamento de engenharia mecânica	(l) estamos no (r) (l implícito - e) estou no (r) (l) estou no (r) (l) estou na frente do (r) (l) to na frente do (r) (l implícito - eu) na frente do (r) (l implícito - eu) estou perto do (r) (l implícito - eu) perto do (r)
Posição A	implícito - eu, eu	gabinetes dos professores	(l implícito - eu) to perto do (r) (l) vejo (r) (l implícito - eu) estou em frente aos (r) (l implícito - eu) to bem a frente (r) (l) vejo (r)
Posição B	salão nobre	segundo andar do prédio da administração	(r) aonde tem (l) (l) no (r)

FONTE: Autora (2021).

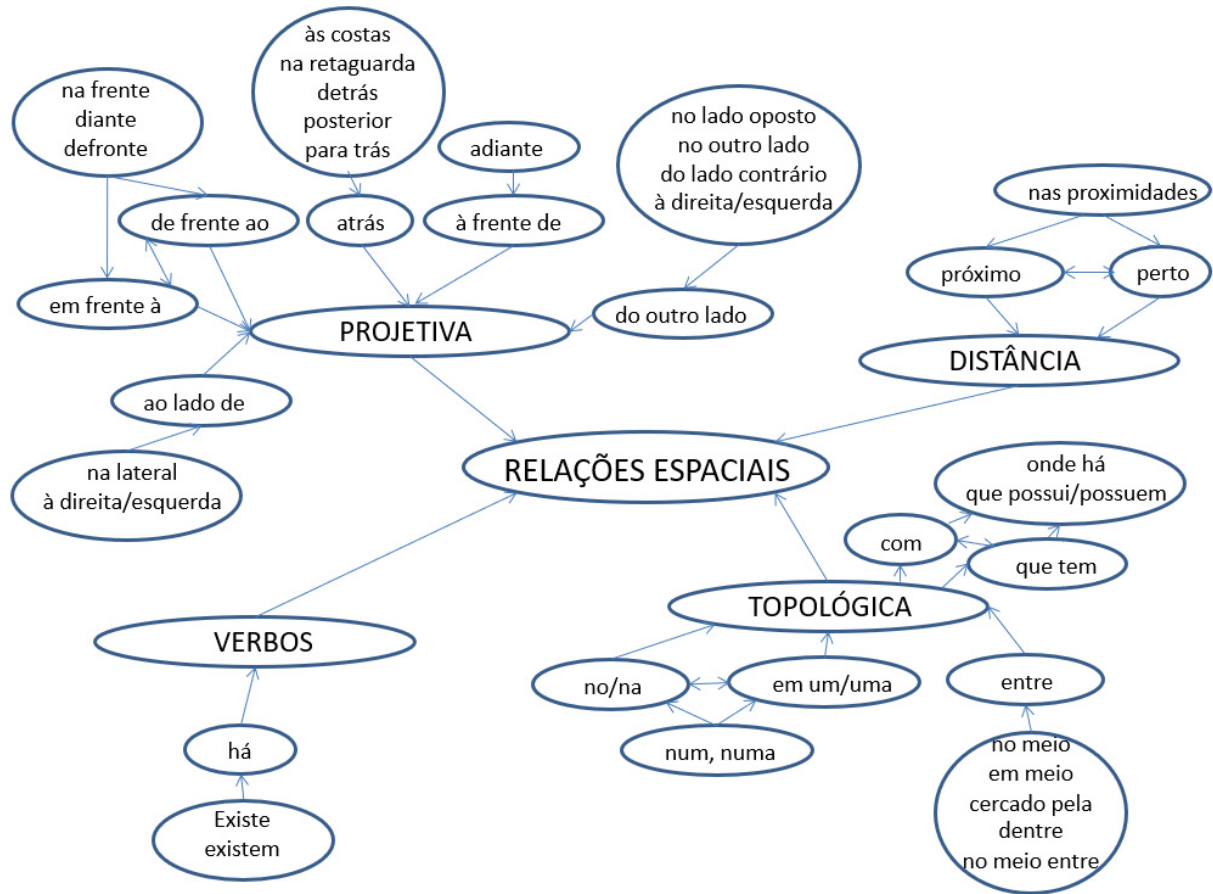
2ª hipótese para o teste de associação livre: Se houver termos de relações espaciais que se comportem como semanticamente similares nas descrições, então pode ser realizado um agrupamento destes, formando uma classe de termos semanticamente similares, conforme a FIGURA 17?

A FIGURA 17 ilustra uma classe de termos semanticamente similares no mesmo nível das classificações existentes de relações espaciais, como por exemplo, topológicas, direcionais e de distância. A hipótese não se verifica porque termos de relações espaciais fazem parte de classes distintas e todos os termos encontrados como semanticamente similares devem pertencer à mesma classe do termo mais citado, ou seja, devem ser topológicos, direcionais, de distância, adjacentes, projetivos, cardinais ou verbos. Isto impossibilita que exista uma classe de termos semanticamente similares no mesmo nível das classes existentes.

Por outro lado, de posse do agrupamento final resultante do teste de “associação livre” pode ser proposta uma rede semântica⁶ de termos de relações espaciais e seus termos semanticamente similares. A rede semântica proposta poderia ter no primeiro nível as classificações das relações espaciais, no segundo nível os termos de relações espaciais mais utilizados e no terceiro nível ficariam os termos encontrados como semanticamente similares, conforme mostra a FIGURA 29. Esta estrutura deve ter as relações “parte-todo”, designadas pelo termo “é parte de”, pois o conhecimento “das partes” de um conceito é útil em compreendê-lo “no todo” (MARCH CASTAÑEDA, 2017). Desta maneira, torna-se possível a análise em forma de rede juntamente com as ligações entre os termos, suas classificações e seus termos semanticamente similares, para aplicações computacionais, principalmente para o PLN e para aplicações de geoprocessamento.

⁶ Rede é uma estrutura de grafo para representar o conhecimento através de nós e arcos interconectados. E as redes semânticas foram desenvolvidas computacionalmente para inteligência artificial e tradução automática (Sowa, 1987).

FIGURA 29 - REDE DE TERMOS DE RELAÇÕES ESPACIAIS



FONTE: A autora (2021).

Para aplicações computacionais, tanto para PLN como para a criação de uma ontologia é necessário realizar a conceituação dos termos, ou seja conceituar e definir as características dos termos de relações espaciais encontrados como similares March Castañeda (2017) explica que uma “categoria” é um conjunto de entidades (instâncias) vinculadas por um princípio unificador, tal como um conjunto de propriedades distintivas, papéis ou funções. E um conceito incorpora todo o conhecimento que descreve o significado de uma categoria. Por exemplo, a categoria “rio” compreende todos os rios do mundo, tais como Amazonas, Nilo, Mississipi, Reno, Amarelo, Eufrates, Congo, enquanto que o conceito “rio” descreve o conhecimento associado àquela categoria, ou seja, o de se referir a um curso de água natural, mais ou menos torrencial, que corre de uma parte mais elevada para uma mais baixa e que deságua em outro rio, no mar ou num lago, que em geral possui meandros e tributários, que rios de grande porte são em geral navegáveis, etc.

A conceituação partiria de uma tabela das características dos termos de relações espaciais com os conceitos, os termos encontrados como semanticamente similares, os elementos que podem fazer parte do *locatum* e do *relatum*, eixo de referência, e quando se deve fazer a utilização do referido termo. Sendo que, os elementos que podem fazer parte tanto dos *locata* quanto dos *relata*, podem ser retirados de ontologias existentes.

A conceituação dos termos, principalmente no que diz respeito às similaridades encontradas deve ser cuidadosamente formulada porque de acordo com Ferreira e Delazari (2019a) o uso indevido de uma palavra para descrever a relação entre elementos pode transmitir informações enganosas sobre a localização dos elementos envolvidos.

Para tanto, a referida tabela (38) deve conter as respostas para as seguintes perguntas: Qual é o conceito de “Em frente”? Quais são os termos semanticamente similares encontrados? Quais elementos podem fazer parte do *locatum* e do *relatum* de expressões locativas com este termo? Qual é o eixo de referência utilizado? Quando este termo de relação espacial pode ser utilizado?

TABELA 38 - CARACTERÍSTICAS DAS RELAÇÕES ESPACIAIS

Termo de relação espacial	Conceito	Termos Semanticamente Similares	Elementos que podem fazer parte do <i>locatum</i> (ontologias)	Elementos que podem fazer parte do <i>relatum</i> (ontologias)	Eixo de referência	Quando utilizar este termo?
Em frente à/às/ao/aos/ de/do/da	Estar na presença de alguém ou de algum animal, feições geográficas, automóveis, elementos de ambientes <i>indoor</i> , elementos de ambientes <i>outdoor</i> .	Na frente De frente Diante De fronte	Eu, nós, a gente, implícito ou explícito	Feições geográficas, pessoas, animais, automóveis, edificações, elementos de ambientes <i>indoor</i> , elementos de ambientes <i>outdoor</i> , elementos pontuais, lineares, de área	Centrado no indivíduo (egocêntrico)	Quando o falante descrever em linguagem natural e (<i>locatum</i>) estiver espacialmente localizado na presença do <i>relatum</i> .
Em frente à/às/ao/aos/ de/do/dos/ da/das	Estar na presença de alguém ou de algum animal, feições geográficas, automóveis, elementos de ambientes <i>indoor</i> , elementos de ambientes <i>outdoor</i> .	Na frente De frente Diante Defronte	Acidentes, assaltos, algo que se queira relatar, feições geográficas, pessoas, animais, automóveis, edificações, elementos de ambientes <i>indoor</i> , elementos de ambientes <i>outdoor</i> , elementos pontuais, lineares, de área	Feições geográficas, pessoas, animais, automóveis, edificações, elementos de ambientes <i>indoor</i> , elementos de ambientes <i>outdoor</i> , elementos pontuais, lineares, de área	Centrado no objeto	Quando o falante descrever em linguagem natural que o <i>locatum</i> está espacialmente localizado na presença do <i>relatum</i> .

FONTE: Autora (2021).

Segundo March Castañeda (2017) convém verificar, com base em dicionários como os termos estudados são entendidos na norma culta da língua portuguesa. E, depois quando da aplicação de uma futura ontologia será necessário investigar como que especialistas empregam os termos estudados.

Porém para a implementação em aplicações computacionais também há a necessidade de saber quais as estruturas de expressões locativas que podem ser aplicadas aos termos de relações espaciais.

4.7 ESTRUTURAÇÃO DE EXPRESSÕES LOCATIVAS COM OS DADOS ADQUIRIDOS NOS TRÊS TESTES

Foram analisadas as expressões locativas dos dois primeiros testes tanto para o eixo egocêntrico como para o eixo centrado nos elementos. A partir da verificação de como os *relata*, os *locata* e os termos de relações espaciais foram utilizados nas expressões locativas, retirou-se então a estrutura de expressões locativas com seus respectivos *relata*, *locata*, termos de relações espaciais que em muitos casos apareceram com vocábulos referentes à gramática da língua portuguesa brasileira para dar sentido às frases. Esta análise resultou em tabelas onde também foram agregados os termos semanticamente similares encontrados no terceiro teste. Por exemplo, para o termo “em frente” tem-se a TABELA 39 para o eixo egocêntrico e 40 para o eixo centrado nos elementos. Os *locata* e os *relata* podem conter termos implícitos.

TABELA 39 – ESTRUTURAS DE EXPRESSÕES LOCATIVAS – TERMO “EM FRENTE” – EIXO EGOCÊNTRICO

Relações Espaciais Semanticamente Similares	Estrutura das frases	Expressões locativas
em/na/de frente a/á/à/às/ao/aos /de/da/das/do/dos	<i>Locatum</i> + RE + <i>Relatum</i> <i>Locatum</i> + verbo "to"/estou/estamos + RE + <i>Relatum</i>	Eu na frente do hospital. Eu estou em frente do museu.
	<i>Locatum</i> + verbo "to"/estou/estamos + verbo localizado/s/situado/s/parado/s+ RE + <i>Relatum</i>	Eu estou localizado em frente à escola. Nós estamos situados em frente de um bosque. Eu estou parado em frente da igreja.
	<i>Locatum</i> + pronomes átonos + verbo encontro/encontramos + RE + <i>Relatum</i>	Eu me encontro em frente ao cinema. Nós nos encontramos em frente à panificadora.
diante a/á/à/às/ao/aos/de/ da/das/do/dos,	verbo encontro-me + pronomes átonos + RE + <i>Relatum</i>	Encontro-me em frente do restaurante.
	<i>Locatum</i> + verbo "to"/estou/estamos + complemento do termo de RE + RE + <i>Relatum</i>	Eu estou bem na frente da casa azul.
defronte a/á/à/às/ao/aos/de/ da/das/do/dos	Verbo tem + <i>Locatum</i> + em/na + pronome possessivo minha/nossa + RE + <i>Relatum</i>	Tem uma banca de jornal na minha/nossa frente.
	<i>Locatum</i> + verbo estar + complemento do termo RE + <i>Relatum</i>	Nós estamos bem na frente das lojas.
	na/em + pronome possessivo minha/nossa + RE + verbo tem	Na nossa frente tem uma praça.
	<i>Locatum</i> + verbo "to"/estou/estamos + advérbio aqui + RE + <i>relatum</i>	Eu to aqui em frente ao colégio.

FONTE: Autora (2021).

Os resultados encontrados nas TABELAS 39 e 40 permitem que as expressões locativas retiradas de descrições em linguagem natural, sejam representadas por “estruturas” de expressões locativas. Percebe-se que pode ser considerado o eixo egocêntrico ou centrado nos elementos, onde os locata podem ser “eu, nós, a gente” explícitos ou implícitos e quaisquer outros elementos/feições geográficas. Os relata também podem ser representados por nomes de entidades ou elementos: City Mary e Starring, e, características de entidades, como: avenida a beira mar, casas residenciais térreas. De acordo com Marcińczuk, Oleksy e Wieczorek (2016), as expressões locativas não obedecem somente à ordem *Locatum* + RE + *Relatum*, elas podem variar de acordo com a gramática da língua analisada.

TABELA 40 – ESTRUTURAS DE EXPRESSÕES LOCATIVAS – TERMO EM FRENTE – EIXO CENTRADO NOS ELEMENTOS

Termos de Relações Espaciais Semanticamente Similares	Estrutura das frases	Expressões locativas
	<i>Locatum + RE + Relatum</i>	Cantina de frente para o corredor principal.
	<i>Locatum + verbo fica/ficam/está/estão/dá/dão + RE + Relatum</i>	Painel fica em frente à porta de saída. Extintor está em frente à Coordenação. Elevador dá de frente pra porta de madeira.
	<i>Locatum + pronome relativo + verbo fica/fica/está/estão/dá/dão + RE + Relatum</i>	Escadas que dão de frente para a saída.
	<i>Locatum + verbo está/estão/fica/ficam/dá/dão + complemento do termo de RE + RE + Relatum</i>	Salão nobre está bem em frente à bilheteria.
em/na/de frente a/à/às/ao/aos/de/das/dos/dos/para/para a/o/as/os/ pra/pro,	<i>RE + Relatum + verbo há/tem/têm/existem/existe + Locatum</i>	Em frente ao banheiro há uma janela. Na frente das escadas tem um extintor. Em frente ao salão nobre existem os banheiros.
	<i>Locatum + complemento do termo de RE bem/quase + RE + Relatum</i>	Elevador bem em frente à recepção.
diante a/à/às/ao/aos/de/das/dos/dos, defronte a/à/às/ao/aos/de/das/dos/dos	<i>Locatum + verbo está/estão + complemento do termo de RE bem/quase + RE + Relatum</i>	Elevador está bem em frente à recepção.
	<i>complemento do termo de RE bem/quase + RE + Relatum + verbo há/tem/têm/existe/existem + Locatum</i>	Quase em frente à porta principal há uma escada.
	<i>Verbo há/tem/têm/existe/existem + Locatum + RE + Relatum</i>	Tem uma porta de vidro na frente do elevador.
	<i>Verbo há/tem/têm/existe/existem + Locatum + complemento do termo de RE bem/quase + RE + Relatum</i>	Há uma mesa bem em frente ao edital.
	<i>Verbo há/tem/têm/existe/existem + advérbio ainda + RE + Relatum + Locatum</i>	Existe ainda em frente à cantina, o departamento de engenharia civil.
	<i>relatum + verbo há/tem/têm/existe/existem + locatum + RE</i>	Os gabinetes dos professores têm um edital na frente.

FONTE: Autora (2021).

5 CONCLUSÕES E RECOMENDAÇÕES

Esta tese teve como objetivo principal a proposição de um método para criação de grupos semanticamente similares de termos de relações espaciais utilizados em linguagem natural do português brasileiro que possam ser utilizados em aplicações computacionais.

Para que a hipótese da pesquisa fosse respondida, foi necessária a realização de três testes, sendo dois testes para coleta de dados em espaços geográficos diferentes e reais, e em dois contextos diferentes (*Indoor* e *Outdoor*), aplicados numa mesma situação através da pergunta: “Onde você está localizado?”; e mais um teste para obtenção dos termos utilizados como similares, que foram confirmados através da aplicação de regras.

Os resultados dos dois primeiros testes e as duas hipóteses propostas para estes dois testes foram fundamentais para o objetivo geral desta pesquisa. Além dos dois testes fornecerem os termos mais utilizados em cada contexto. Pode-se afirmar que na língua portuguesa brasileira existe um grupo de termos de relações espaciais (em frente de/da/do, no/na/em, próximo, ao lado de, lado direito/esquerdo e perto) que variam pouco para descrições de ambientes com diferentes configurações e contextos. Portanto, além de existir um grupo de termos que são utilizados comumente no dia a dia, quando fazem suas descrições as pessoas os utilizam tanto no referencial egocêntrico, quanto no centrado nos elementos.

A hipótese de que os termos de relações espaciais não se apresentaram dependentes do contexto de uso, traz a descoberta de que para a língua portuguesa falada/escrita no Brasil os termos de relações espaciais mais utilizados podem ser explorados na sua totalidade para diversos contextos diferentes, como por exemplo, urbanos, rurais, internos, externos, em situações simuladas ou do cotidiano das pessoas. Isto contribui para implementações de aplicações em SIG, considerando a formulação/interpretação de perguntas para transformar as descrições em linguagem natural em representações gráficas; e aumentando o grupo das relações espaciais que podem ser utilizadas nas funções de geoprocessamento.

Os resultados do terceiro teste trouxeram não só os termos utilizados, como também, possibilitaram a aplicação das regras que foram elaboradas para facilitar nas análises para a elaboração dos agrupamentos finais dos termos que são semanticamente similares. As duas hipóteses criadas para análise dos termos

encontrados também foram de grande valia, mostrando que existiram termos que foram utilizados nos dois primeiros testes que já se apresentavam como similares; e que não é possível criar uma única classe de termos semanticamente similares pensando-se numa rede semântica, porque os termos pertencem a classificações de relações espaciais diferentes. Esta tese também apresenta a classificação dos termos de acordo com as classes já existentes na literatura e não somente a classificação topológica que é a mais explorada.

As regras elaboradas funcionaram como um filtro facilitando a análise final do agrupamento, e a classificação dos termos de relações espaciais facilita a utilização destes termos em aplicações computacionais voltadas para o PLN e para a criação de ontologias.

A contabilização da quantidade de citações permitiu a análise dos termos que são mais similares entre si pelo seu distanciamento decrescente. Isto facilita a utilização dos termos em aplicações computacionais que necessitem a utilização de mais de um termo, como por exemplo, a geração de descrições de localização em linguagem natural. Além disso, os termos encontrados podem ser utilizados no conjunto de interfaces do projeto *Where am I?* respeitando a classificação de cada termo. Portanto, os termos encontrados podem ser utilizados em aplicações computacionais, tanto para PLN como em colaboração com programas de SIG.

Recomenda-se como trabalhos futuros a realização de testes práticos com a estruturação das expressões locativas encontradas neste estudo.

Sugere-se aplicar a metodologia dos testes utilizando a abordagem presencial para o ambiente *outdoor* e a metodologia *online* para o ambiente *indoor*, visto que os resultados mostraram que 86% dos participantes utilizaram o eixo centrado nos elementos no teste *outdoor* (realizado *online*) e 59% o eixo centrado no próprio observador no *indoor* (presencial). O ambiente *indoor* escolhido poderia ser um *shopping center* classificado como um ambiente comercial.

Além disso, como os dados foram coletados por nativos da língua portuguesa brasileira tem-se a possibilidade de aplicar este experimento para nativos da língua inglesa e fazer uma análise interlinguística entre os termos encontrados nas duas línguas, pois pode ser investigado o grau de impacto das diferenças das duas línguas nos métodos automatizados de georreferenciamento e geração de linguagem espacial entre o inglês e o português brasileiro. Pois, estudos comparando as duas línguas são escassos, e os estudos interlinguísticos são

frequentemente em ambientes controlados e apresentam aspectos específicos somente de alguns termos. Estudos com situações do mundo real, em espaços geográficos reais introduzem oportunidades adicionais para estudar diferenças na seleção de características geográficas particulares e nos termos de relação espacial. Isso contribuiria com a correta tradução de termos de relação espacial para ser aplicada computacionalmente. Também seria possível fazer uma comparação dos termos apresentados numa enciclopédia multilíngue como a *BabelNet* que apresenta traduções e redes semânticas em diversas línguas.

Futuramente também se devem aplicar mais testes nos termos encontrados com sugestão de substituição de relações espaciais para validação das similaridades encontradas. E a realização de uma ontologia para relações espaciais.

REFERÊNCIAS

- ALMEIDA, M. B.; SOUZA, R. R. Avaliação do espectro semântico de instrumentos para organização da informação. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 16, n. 31, p. 25–50, 25 maio 2011.
- ARAÚJO, P. J. P. **Aspectos semântico-cognitivos de usos espaciais das preposições para e em na fala de comunidades quilombolas**. São Paulo: Universidade de São Paulo, 2008.
- BATEMAN, J. A.; HOIS, J.; ROSS, R.; TENBRINK, T. A linguistic ontology of space for natural language processing. **Artificial Intelligence**, v. 174, n. 14, p. 1027–1071, set. 2010.
- BATORÉO, H. J. **Language typology and semantic primitive of space: evidence from European Portuguese**. (M. Vilela, F. Silva, Eds.) Actas do 1º Encontro Internacional de Linguística Cognitiva. **Anais...**: Actas do 1º Encontro Internacional de Linguística Cognitiva. In: 1º ENCONTRO INTERNACIONAL DE LINGUÍSTICA COGNITIVA. Porto: Faculdade de Letras do Porto, 1998.
- BATORÉO, H. J.; DUARTE, I. **PARA A CARACTERIZAÇÃO DA LÍNGUA-ALVO NO ENSINO DE PORTUGUÊS COMO LÍNGUA ESTRANGEIRA: UM ESTUDO DE DESENVOLVIMENTO LINGUÍSTICO DE ABERTURAS DE NARRATIVAS**. Actas do V Congresso Internacional para o Ensino de Português Como Língua Estrangeira. **Anais...** In: V CONGRESSO INTERNACIONAL PARA O ENSINO DE PORTUGUÊS COMO LÍNGUA ESTRANGEIRA. Cidade do México: 2001.
- BELOUAER, L.; CLARAMUNT, C. **Modeling Spatial Knowledge from Verbal Descriptions**. Spatial Information Theory. **Anais...** In: 11TH INTERNATIONAL CONFERENCE, COSIT 2013. Scarborough, UK: 2013.
- BENNETT, B.; AGARWAL, P. Semantic Categories Underlying the Meaning of 'Place'. In: WINTER, S. et al. (Eds.). **Spatial Information Theory**. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007. v. 4736p. 78–95.
- BERTHELE, R.; WHELPTON, M.; NAESS, A.; DUIJFF, P. Static spatial descriptions in five Germanic languages. **Language Sciences**, v. 49, p. 82–101, maio 2015.
- BERTHELE, R. Convergence in the domains of static spatial relations and events of putting and taking. Evidence from bilingual speakers of Romansh and German. **International Journal of Bilingual Education and Bilingualism**, v. 18, n. 5, p. 624–642, 3 set. 2015.
- BITTERS, B. Spatial Relationship Networks: Network Theory Applied to GIS Data. **Cartography and Geographic Information Science**, v. 36, n. 1, p. 81–93, 1 jan. 2009.
- BLAYLOCK, N.; SWAIN, B.; ALLEN, J. Mining Geospatial Path Data from Natural Language Descriptions. In: **Proc. ACM SIGSPATIAL GIS International Workshop on Querying and Mining Uncertain Spatio-Temporal Data.**, p. 6, 2009.

BLÜHDORN, H. **A Codificação de Informação Espacial no Alemão e no Português do Brasil**. [s.l.] Universidade de São Paulo, 2001.

BOSSE, S.; PAPAFRAGOU, A. Spatial Position in Language and Visual Memory: A Cross-Linguistic Comparison. In: **Proceedings of the Annual Meeting of the Cognitive Science Society**., v. 32, n. 32, p. 7, 2010.

BOWERMAN, M.

Learning How to Structure Space for Language: A Crosslinguistic Perspective. In: BLOOM, P. et al. (Eds.). **Language and Space**. [s.l.] The MIT Press, 1996. p. 385–436.

BREITMAN, K. K. Ontologias – Como e Porquê Criá-las. p. 50, 2004.

BROWN, P. Up, down, and across the land: landscape terms, place names, and spatial language in Tzeltal. **Language Sciences**, v. 30, n. 2–3, p. 151–181, mar. 2008.

BRUNS, H. T.; EGENHOFER, M. J. Similarity of Spatial Scenes. In: **Seventh international symposium on spatial data handling**., n. Delft, The Netherlands, p. 31–42, 1996.

BUNGER, A.; SKORDOS, D.; TRUESWELL J. C.; PAPAFRAGOU, A. How children and adults encode causative events cross-linguistically: implications for language production and attention. **Language, Cognition and Neuroscience**, v. 31, n. 8, p. 1015–1037, 13 set. 2016.

CAMBRIA, E.; WHITE, B. Jumping NLP Curves: A Review of Natural Language Processing Research [Review Article]. **IEEE Computational Intelligence Magazine**, v. 9, n. 2, p. 48–57, maio 2014.

CARLSON, L. A.; LOGAN, G. D. Using spatial terms to select an object. **Memory & Cognition**, v. 29, n. 6, p. 883–892, 2001.

CARLSON-RADVANSKY, L. A.; LOGAN, G. D. The Influence of Reference Frame Selection on Spatial Template Construction. **Journal of Memory and Language**, v. 37, n. 3, p. 411–437, out. 1997.

CHEN, H.; VASARDANI, M.; WINTER, S.; TOMKO, M. A Graph Database Model for Knowledge Extracted from Place Descriptions. **ISPRS International Journal of Geo-Information**, v. 7, n. 6, p. 221, 15 jun. 2018.

CLARK, H. H. SPACE, TIME, SEMANTICS, AND THE CHILD. **Cognitive development and acquisition of language**, Academic Press. p. 27–63, 1973.

CLEMENTINI, E. A Conceptual Framework for Modelling Spatial Relations. **Information Technology And Control**, v. 48, n. 1, p. 5–17, 25 mar. 2019.

CLEMENTINI, E.; DI FELICE, P. Approximate topological relations. **International Journal of Approximate Reasoning**, v. 16, n. 2, p. 173–204, fev. 1997.

CLEMENTINI, E.; SHARMA, J.; EGENHOFER, M. J. Modelling topological spatial relations: Strategies for query processing. **Computers & Graphics**, v. 18, n. 6, p. 815–822, 1 nov. 1994.

COHN, A. G.; BENNET, B.; GOODAY, J.; GOTTS, N. M. Qualitative Spatial Representation and Reasoning with the Region Connection Calculus. **Geoinformatica**, v. 1, n. 3, p. 275–316, 1 out. 1997.

COUTO, H. H. DO. THE ECOLOGY OF SPATIAL RELATIONS: THE CASE OF KRIOL PREPOSITIONS. In: SCHRADER-KNIFFKI, M.; GARCÍA, L. M. (Eds.). **La Romania en interacción**. [s.l.] Vervuert Verlagsgesellschaft, 2007. p. 479–514.

COVENTRY, K.; GUIJARRO-FUENTES, P.; VALDÉS, B. On the First and Second Language Acquisition of Spatial Language. **Spatial Cognition & Computation**, v. 12, n. 4, p. 219–230, out. 2012.

COVENTRY, K. R. On the mapping between spatial language and the vision and action systems. In: **Language and Action in Cognitive Neuroscience**. [s.l.: s.n.]. p. 209–225, 2013.

COVENTRY, K. R.; GARROD, S. C. **Saying, Seeing and Acting: The Psychological Semantics of Spatial Prepositions**. [s.l.] Psychology Press, 2004.

COYNE, B.; SPROAT, R.; HIRSCHBERG, J. Spatial Relations in Text-to-Scene Conversion. In: **Computational Models of Spatial Language Interpretation, Workshop at Spatial Cognition**, p. 8, 2010.

CUAYÁHUITL, H.; DETHLEFS, N.; RISHTER, K.; TENBRINK, T.; BATEMAN, J. A Dialogue System for Indoor Wayfinding Using Text-Based Natural Language. **International Journal of Computational Linguistics and Applications**, v. 1, n. 1-2, p. 285-304, 2010., v. 1, n. 1–2, p. 285–304, 2010.

DE BRITO, M. Sistemas de informação em linguagem natural: em busca de uma indexação automática. **Ciência da Informação**, v. 21, n. 3, p. 10, 1992.

DE FREITAS, F. L. G. Ontologias e a web semântica. **Jornada de Mini-Cursos em Inteligência Artificial - SBC**, v. 8, 2003.

DELAZARI, L. S.; STOCK, K.; FERREIRA, M. E. DOS S.; FAGUNDES, C. K. M. **Use of Spatial Relations in English and Portuguese: Syntax and Semantics**. ICC Conference. **Anais...** In: 28TH INTERNATIONAL CARTOGRAPHIC CONFERENCE. Washington: 2017

DENIS, M. The Description of Routes: A Cognitive Approach to the Production of Spatial Discourse. **Current Psychology of Cognition**, v. 16, n. 4, p. 409–458, 1997.

DERUNGS, C.; PURVES, R. S. Mining nearness relations from an n-grams Web corpus in geographical space. **Spatial Cognition & Computation**, v. 16, n. 4, p. 301–322, out. 2016.

EGENHOFER, M. J.; CLEMENTINI, E.; DI FELICE, P. Topological Relations between Regions with Holes. **International Journal of Geographical Information Systems**, v. 8, n. 2, p. 129–144, 1994.

EGENHOFER, M. J.; HERRING, J. R. Categorizing Binary Topological Relations Between Regions, Lines, and Points in Geographic Databases. **The**, . The, v. 9, n. 94-1, p. 76, 1990. v. 9, n. 94–1, p. 28, 1990.

EGOROVA, E. **Spatial Discourse Production: Applying Denis’s Framework to Non-urban Context**. (S. Creem-Regehr, J. Schöning, A. Klippel, Eds.) Spatial Cognition XI. **Anais...**: Lecture Notes in Computer Science. In: 11TH INTERNATIONAL CONFERENCE, SPATIAL COGNITION. Tübingen, Germany: Springer International Publishing, 2018. Disponível em: <<http://link.springer.com/10.1007/978-3-319-96385-3>>. Acesso em: 5 mar. 2021

FAGUNDES, C. K. M; DELAZARI, L. S. Relações Espaciais Obtidas a Partir de Descrições de Localização em Linguagem Natural em Ambientes Indoor e Outdoor. **Anuário do Instituto de Geociências**. v.44, DOI 10.11137/1982-3908_2021_44_36896.

FANG, Z.; LI, Q.; SHAW, S.-L. What about people in pedestrian navigation? **Geo-spatial Information Science**, v. 18, n. 4, p. 135–150, 2 out. 2015.

FEIST, M. I. Talking about space: A cross-linguistic perspective. **Proceedings of the Annual Meeting of the Cognitive Science Society**, v. 26, p. 375–380, 2004.

FERREIRA, A. B. DE H. **Dicionário Aurélio da língua portuguesa**. [s.l.] Portal UOL, 2010.

FERREIRA, M.; CLARO, D. B.; LOPES, D. Integração do Mecanismo de Self-healing na Execução das Composição de Sistemas de Informação através dos Serviços Web Semânticos. In: **VII Simpósio Brasileiro de Sistemas de Informação**., p. 12, 2011.

FERREIRA, M. E. DOS S. **PROPOSIÇÃO DE UM CONJUNTO DE RELAÇÕES ESPACIAIS PARA TAREFAS DE DESCRIÇÕES ESPACIAIS**. Tese—Curitiba: UNIVERSIDADE FEDERAL DO PARANÁ, 2019.

FERREIRA, M. E. DOS S.; DELAZARI, L. S. THE USE OF SPATIAL TERMS “NEAR”, “VERY NEAR”, “NEXT TO”, “SIDE BY SIDE AND “NEARBY” IN THE DESCRIPTIONS OF SPATIAL CONFIGURATIONS. **Boletim de Ciências Geodésicas**, v. 25, n. 2, p. e2019008, 2019a.

FERREIRA, M. E. DOS S.; DELAZARI, L. S. Using Spatial Image Schemata in the characterization of spatial relations. **Proceedings of the International Cartographic Association**, v. 2, p. 1–6, 10 jul. 2019b.

FILLMORE, CH. J. Santa Cruz Lectures on Deixis, Indiana University Linguistics Club, Bloomington. (1975 [1971]).

FRANCO, L.; ZAMPIERI, E.; MENEGHELLO, F. Prepositions inside (and at the edge) of words: a view from agrammatism. **Language Sciences**, v. 40, p. 95–122, nov. 2013.

FREITAS, J. B. **SiSe: Medida de similaridade semântica entre ontologias em português**. Dissertação—Rio Grande do Sul: Pontifícia Universidade Católica do Rio Grande do Sul, ago. 2007.

GENTNER, D.; OZYUREK, A.; GURCANLI, O.; GOLDIN-MEADOW, S. Spatial language facilitates spatial cognition: Evidence from children who lack language input. **Cognition**, v. 127, n. 3, p. 318–330, jun. 2013.

GRIGOROGLOU, M.; PAPAFRAGOU, A. Spatial Terms. In: **Handbook of Experimental Semantics and Pragmatics**. [s.l.] Oxford University Press, 2018.

GRUBER, T. R. A translation approach to portable ontology specifications. **Knowledge Acquisition**, v. 5, n. 2, p. 199–220, jun. 1993.

GUY, G. R. **LINGUISTIC VARIATION IN BRAZILIAN PORTUGUESE: ASPECTS OF THE PHONOLOGY, SYNTAX, AND LANGUAGE HISTORY**. PhD Thesis—[s.l.] University of Pennsylvania, 1981.

HALL, M. M.; JONES, C. B. Quantifying Spatial Prepositions: An Experimental Study. In: **Proceedings of the 16th ACM SIGSPATIAL international conference on advances in geographic information systems.**, p. 1–4, 2008.

HAMZEI, E.; WINTER, S.; TOMKO, M. **Initial Analysis of Simple Where-Questions and Human-Generated Answers (Short Paper)**. In: 14th International Conference on Spatial Information Theory. **Anais...** In: COSIT 2019. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik: Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik GmbH, Wadern/Saarbruecken, Germany, 2019. Disponível em: <<http://drops.dagstuhl.de/opus/volltexte/2019/11104/>>. Acesso em: 5 mar. 2021

HAN, L.; KASHYAP, A.; FININ, T.; MAYFIELD, J.; WEESE, J. **UMBCEBIQUITY-CORE: Semantic Textual Similarity Systems**. Proceedings of the Main Conference and the Shared Task. **Anais...** In: SECOND JOINT CONFERENCE ON LEXICAL AND COMPUTATIONAL SEMANTICS. Atlanta, Georgia: jun. 2013

HATZIVASSILOGLOU, V.; WIEBE, J. M. **Effects of adjective orientation and gradability on sentence subjectivity**. Proceedings of the 18th conference on Computational linguistics -. **Anais...** In: THE 18TH CONFERENCE. Saarbrücken, Germany: Association for Computational Linguistics, 2000. Disponível em: <<http://portal.acm.org/citation.cfm?doid=990820.990864>>. Acesso em: 3 mar. 2021

HERNÁNDEZ-PASTOR, D.; PERIÑÁN-PASCUAL, C.; VALENCIA, U. P. DE. Developing a knowledge base for preposition sense disambiguation: A view from Role and Reference Grammar and FunGramKB. **Onomázein Revista de lingüística, filología y traducción**, n. 33, p. 251–288, 5 ago. 2016.

HERSKOVITS, A. Semantics and pragmatics of locative expressions. **Cognitive Science**, v. 9, n. 3, p. 341–378, 1 jul. 1985.

HIMMELBERGER, Z. M. **Processing spatial relations: the role of instructions on the priming of egocentric and allocentric spatial representations**. thesis—[s.l.] University of Alabama Libraries, 2015.

HOFFMANN, D. Restrictions on the Usage of Spatial Frames of Reference in Location and Orientation Descriptions: Evidence from Three Australian Languages. **Australian Journal of Linguistics**, v. 39, n. 1, p. 1–31, 2 jan. 2019.

HOIS, J.; TENBRINK, T.; ROSS, R.; BATEMAN, J. **GUM-Space: The Generalized Upper Model spatial extension: a linguistically-motivated ontology for the semantics of spatial language**: SFB/TR8 Spatial Cognition. [s.l.] Universität Bremen, 2009. Disponível em: <<http://www.ontospace.uni-bremen.de/ontology/TechnReport09GUMspace.pdf>>.

HU, Y. Geo-text data and data-driven geospatial semantics. **Geography Compass**, v. 12, n. 11, nov. 2018.

JACKMAN, H. J. D. SEMANTIC NORMS AND TEMPORAL EXTERNALISM. p. 242, 1996.

JUNGBLUTH, K. Os pronomes demonstrativos do Português Brasileiro na fala e na escrita. **Cadernos de Linguagem e Sociedade**, v. 7, p. 83–105, 2005.

KEMMERER, D. “Near” and “far” in language and perception. **Cognition**, v. 73, n. 1, p. 35–63, nov. 1999.

KEMMERER, D. The semantics of space: Integrating linguistic typology and cognitive neuroscience. **Neuropsychologia**, New Insights in Categorical and Coordinate Processing of Spatial Relations. v. 44, n. 9, p. 1607–1621, 1 jan. 2006.

KEWITZ, V. A representação do movimento no português paulista. **Filologia e Linguística Portuguesa**, v. 13, n. 1, p. 89, 4 jun. 2011.

KLATZKY, R. L. Allocentric and Egocentric Spatial Representations: Definitions, Distinctions, and Interconnections. **Spatial cognition - An interdisciplinary approach to representation and processing of spatial knowledge**, p. 1–17, 1998.

KLEPPA, L.-A. PREPOSIÇÕES MAIS GRAMATICALIZADAS EM DICIONÁRIOS ESCOLARES. **Veredas-Revista de Estudos Linguísticos**, v. 12, n. 1, p. 17, 2008.

KORDJAMSHIDI, P.; FRASCONI, P.; OTTERLO, M. V.; MOENS, M.-F.; RAEDT, L. Relational Learning for Spatial Relation Extraction from Natural Language. In: MUGGLETON, S. H.; TAMADDONI-NEZHAD, A.; LISI, F. A. (Eds.). **Inductive Logic Programming**. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. v. 7207p. 204–220.

KORDJAMSHIDI, P.; HOIS, J.; OTTERLO, M. V.; MOENS, M.-F. Learning to interpret spatial natural language in terms of qualitative spatial relations*. In: TENBRINK, T.; WIENER, J. M.; CLARAMUNT, C. (Eds.). **Representing Space in Cognition**. [s.l.] Oxford University Press, 2013. p. 115–146.

KORDJAMSHIDI, P.; OTTERLO, M. V.; MOENS, M.-F. From Language towards Formal Spatial Calculi. In: **Proc. of 1st Workshop COSLI'10**, v. 620, p. 17–24, 2010.

KREMER, G.; BARONI, M. A set of semantic norms for German and Italian. **Behavior Research Methods**, v. 43, n. 1, p. 97–109, mar. 2011.

KUHN, W. Ontologies in support of activities in geographical space. **International Journal of Geographical Information Science**, v. 15, n. 7, p. 613–631, out. 2001.

KURATA, Y.; SHI, H. Interpreting Motion Expressions in Route Instructions Using Two Projection-Based Spatial Models. In: DENGEL, A. R. et al. (Eds.). **KI 2008: Advances in Artificial Intelligence**. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. v. 5243p. 258–266.

LAKOFF, G. *Women, fire and dangerous things. What categories reveal about the mind*, University of Chicago Press, Chicago. 1987.

LANDAU, B.; JACKENDOFF, R. Whence and whither in spatial language and spatial cognition? **Behavioral and Brain Sciences**, v. 16, n. 2, p. 255–265, jun. 1993.

LANGACKER, R. W. *Foundations of cognitive grammar. Vol. I: Theoretical Prerequisites*, Stanford University Press, Stanford, CA, 1987.

LEVINSON, S.; MEIRA, S.; GROUP, T. L. AND C. “Natural Concepts” in the Spatial Topological Domain-Adpositional Meanings in Crosslinguistic Perspective: An Exercise in Semantic Typology. **Language**, v. 79, n. 3, p. 485–516, 2003.

LI, R.; HUANG, H. Staying oriented in wayfinding: a comparison of orientation information between english and mandarin speakers. **GeoJournal**, 13 jun. 2020.

LIPSKI, J. M. External history and linguistic change: Brazilian Portuguese-s. **Luso-Brazilian Review**, v. 12, n. 2, p. 213–224, 1975.

LIU, F.; VASARDANI, M.; BALDWIN, T. **Automatic Identification of Locative Expressions from Social Media Text: A Comparative Analysis**. Proceedings of the 4th International Workshop on Location and the Web - LocWeb '14. **Anais...** In: THE 4TH INTERNATIONAL WORKSHOP. Shanghai, China: ACM Press, 2014Disponível em: <<http://dl.acm.org/citation.cfm?doid=2663713.2664426>>. Acesso em: 26 dez. 2020

LOPES, I. L. Uso das linguagens controlada e natural em bases de dados: revisão da literatura. **Ciência da Informação**, v. 31, n. 1, p. 41–52, jan. 2002.

LYNCH, K. **The image of the city**. Cambridge: Massachusetts Institute of Technology. Press., 1960.

MACHADO, P. C. A INFLUÊNCIA DA LINGUAGEM VISO-ESPACIAL NO DESENVOLVIMENTO COGNITIVO DA CRIANÇA SURDA. **Revista Linhas**, v. 2, n. 2, p. 14, 2001.

MAJID, A.; JORDAN, F.; DUNN, M. Semantic systems in closely related languages. **Language Sciences**, v. 49, p. 1–18, maio 2015.

MARCH CASTAÑEDA, R. **UMA ONTOLOGIA DE REFERÊNCIA PARA O DOMÍNIO DAS DIVISÕES E LIMITES GEOGRÁFICOS TERRITORIAIS DO BRASIL**. Tese - volume I e II—Rio de Janeiro: UNIVERSIDADE FEDERAL DO RIO DE JANEIRO, 2017.

MARCIŃCZUK, M. M.; OLEKSY, M.; WIECZOREK, J. Towards Recognition of Spatial Relations between Entities for Polish. **Cognitive Studies | Études cognitives**, n. 16, p. 119–132, 31 dez. 2016.

MARK, D. M.; COMAS, D.; EGENHOFER, M. J.; FREUNDSCHUH, S. M.; GOULD, M. D.; NUNES, J. Evaluating and refining computational models of spatial relations through cross-linguistic human-subjects testing. In: FRANK, A. U.; KUHN, W. (Eds.). **Spatial Information Theory A Theoretical Basis for GIS**. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 1995. v. 988p. 553–568.

MARK, D. M.; EGENHOFER, M. J. **Topology of Prototypical Spatial Relations Between Lines and Regions in English and Spanish**. In: Proceedings of the Twelfth International Symposium on Computer- Assisted Cartography. Volume 4. **Anais...**1995

MARK, D. M.; GOULD, M. D.; NUNES, J. SPATIAL LANGUAGE AND GEOGRAPHIC INFORMATION SYSTEMS: CROSS-LINGUISTIC ISSUES (EL LENGUAJE ESPACIAL y LOS SISTEMAS DE INFORMACION GEOGRAFICOS. **Temas Interlinguísticos, Proceedings II Conferencia Latinoamericana sobre el Tecnologia de los Sistemas de Informacion Geograficos**, 1989.

MAROTTA, G. (ED.). **Space in language: proceedings of the Pisa International Conference**. Pisa: ETS, 2010.

MATSUSHIMA, E. H.; RIBEIRO FILHO, N. P. INTERAÇÕES ENTRE SISTEMAS DE REFERÊNCIA ALOCÊNTRICOS E EGOCÊNTRICOS: EVIDÊNCIAS DOS ESTUDOS COM DIREÇÃO PERCEBIDA. **Estudos e Pesquisas em Psicologia**, v. 3, n. 1, p. 72–81, 2003.

MCCLEARY, L.; VIOTTI, E. Representação do espaço em inglês e português brasileiro: observações iniciais. **Revista da Anpoll**, v. 1, n. 16, 15 dez. 2004.

MCDONOUGH, L.; CHOI, S.; MANDLER, J. M. Understanding spatial relations: Flexible infants, lexical adults. **Cognitive Psychology**, v. 46, n. 3, p. 229–259, maio 2003.

MCNAMARA, T. P.; RUMP, B.; WERNER, S. Egocentric and geocentric frames of reference in memory of large-scale space. **Psychonomic Bulletin & Review**, v. 10, n. 3, p. 589–595, set. 2003.

Michaelis Moderno Dicionário Inglês & Português. . [s.l.] Editora Melhoramentos Ltda., 2021.

MOLDOVAN, C. D.; FERRE, P.; DEMESTRE, J.; SANCHEZ-CASAS, R. Semantic similarity: normative ratings for 185 Spanish noun triplets. **Behavior Research Methods**, v. 47, n. 3, p. 788–799, set. 2015.

MONTELLO, D. R. Scale and multiple psychologies of space. In: FRANK, A. U.; CAMPARI, I. (Eds.). **Spatial Information Theory A Theoretical Basis for GIS**. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 1993. v. 716p. 312–321.

MORAIS, E. A. M.; AMBRÓSIO, A. P. L. **Ontologias: conceitos, usos, tipos, metodologias, ferramentas e linguagens**: Instituto de Informática. [s.l.] Universidade Federal de Goiás, 2007.

MUNNICH, E.; LANDAU, B.; DOSHER, B. A. Spatial language and spatial representation: a cross-linguistic comparison. **Cognition**, v. 81, n. 3, p. 171–208, out. 2001.

NOORDZIJ, M. L.; ZUIDHOEK, S.; POSTMA, A. The influence of visual experience on the ability to form spatial mental models based on route and survey descriptions. **Cognition**, v. 100, n. 2, p. 321–342, jun. 2006.

NOTH, W. **Handbook of semiotics**. Bloomington: Indiana University Press, 1990.

NOY, N. F.; MCGUINNESS, D. L. Ontology Development 101: A Guide to Creating Your First Ontology. p. 25, 2001.

OGDEN, C. K.; RICHARDS, I. A. **The Meaning of Meaning: a study of the influence of language upon thought and of the science of symbolism**. New York: A Harvest Book, 1923.

OLIVEIRA, A. DE A. Cognitive relations in the semantics of Brazilian-Portuguese preposition em. **The UK Cognitive Linguistics Association**, v. 1, p. 19–45, 2012.

OLIVEIRA, R. G. **Applying the General Upper Model for Automatic Generation of Spatial Language in Brazilian Portuguese** 3-oliveira-ma.thesis.rgdo.pdf. [s.l.] Bremen, 2013.

PALMER, B.; BLYTHE, J.; GABY, A.; HOFFMANN, D.; PONSONNET, M. **Geospatial Natural Language in Indigenous Australia: Research Priorities**. In: Proceedings Speaking of Location 2019: Communicating about Space. **Anais...**Regensburg, Germany: 2019

PANCHENKO, A. **Similarity Measures for Semantic Relation Extraction**. Tese—Louvain-la Neuve: Université catholique de Louvain & Bauman Moscow State Technical University, 2013.

PAPAFRAGOU, A.; MASSEY, C.; GLEITMAN, L. Shake, rattle, ‘n’ roll: the representation of motion in language and cognition. **Cognition**, v. 84, n. 2, p. 189–219, jun. 2002.

PEREIRA, M. C. J. Aspectos Semânticos e Pragmáticos de aqui, aí, ali, cá e lá em Português Europeu Mário Crescêncio de Jesus Pereira Aspectos Semânticos e Pragmáticos de aqui, aí, ali, cá e lá. **Dissertação de Mestrado**, p. 117 pp, 2009.

PESQUITA, C.; FARIA, D.; FALCÃO, A. O.; LORD, P.; COUTO, F. M. Semantic Similarity in Biomedical Ontologies. **PLoS Computational Biology**, v. 5, n. 7, p. e1000443, 31 jul. 2009.

PINKER, S.; RAVAGNANI, F. **Do que é feito o pensamento: a língua como janela para a natureza humana**. Editora Companhia das Letras. [s.l.] Editora Companhia das Letras, 2008.

RAGNI, M.; TSEDEN, B.; KNAUFF, M. **Cross-Cultural Similarities in Topological Reasoning**. (S. Winter et al., Eds.) Spatial Information Theory. **Anais...: Lecture Notes in Computer Science**. Berlin, Heidelberg: Springer, 2007

REGIER, T. P. The acquisition of lexical semantics for spatial terms: A connectionist model of perceptual categorization. **University of California, Berkeley**, v. 92, n. 62, p. 198, 1992.

REGIER, T.; ZHENG, M. Attention to Endpoints: A Cross-Linguistic Constraint on Spatial Meaning. **Cognitive Science**, v. 31, n. 4, p. 705–719, 8 jul. 2007.

RETZ-SCHMIDT, G. Various Views on Spatial Prepositions. **AI Magazine**, v. 9, n. 2, p. 95–95, 15 jun. 1988.

RHEE, S. Semantic Structure of English Prepositions: An Analysis from a Grammaticalization Perspective. **Language Research**, v. 40, n. 2, p. 397–427, 2004.

RICHTER, K.-F.; WINTER, S. **Landmarks: GIScience for Intelligent Services**. 1st ed. 2014 ed. Cham: Springer International Publishing: Imprint: Springer, 2014. v. 10

RICHTER, K.-F.; WINTER, S.; RÜETSCHI, U.-J. **Constructing Hierarchical Representations of Indoor Spaces**. 2009 Tenth International Conference on Mobile Data Management: Systems, Services and Middleware. **Anais...Taipei, Taiwan: IEEE**, 2009. Disponível em: <<http://ieeexplore.ieee.org/document/5089024/>>. Acesso em: 5 mar. 2021

RODRIGUES, E.; BENNETT, B. **Formalism for Treatment of the Ambiguity in Front/Back Axis Expressions**. [s.l.: s.n.] 2017.

RODRIGUES, E.; GARCIA, D. **A formalism for the treatment of polysemy of vertical prepositions in Brazilian Portuguese**. The First Winter AAAI. **Anais... In: TWENTY-NINTH AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE (AAAI-15)**. Austin Texas, USA: 2015

RODRIGUES, E.; SANTOS, P. E.; LOPES, M. Pinning down polysemy: A formalisation for a Brazilian Portuguese preposition. **Cognitive Systems Research**, v. 41, p. 84–92, mar. 2017.

RODRIGUEZ, M. A.; EGENHOFER, M. J. Determining semantic similarity among entity classes from different ontologies. **IEEE Transactions on Knowledge and Data Engineering**, v. 15, n. 2, p. 442–456, mar. 2003.

SABINO, A. R.; HEINZLE, R. Ferramenta para Construção de Ontologia a Partir de Dados Não Estruturados. **Anais do Computer on the Beach**, p. 141–150, 2015.

SALES, L. F.; CAMPOS, M. L. DE A.; GOMES, H. E. Ontologias de domínio: um estudo das relações conceituais. **Perspectivas em Ciência da Informação**, v. 13, n. 2, p. 62–76, 2008.

SCHMID, F.; KUNTZSCH, C.; WINTER, S.; KAZERANI, A.; PREISIG, B. **Situated local and global orientation in mobile you-are-here maps**. Proceedings of the 12th international conference on Human computer interaction with mobile devices and services - MobileHCI '10. **Anais...** In: THE 12TH INTERNATIONAL CONFERENCE. Lisbon, Portugal: ACM Press, 2010Disponível em: <<http://portal.acm.org/citation.cfm?doid=1851600.1851617>>. Acesso em: 4 mar. 2021

SCHMIDT, M. A. R. **USO DE MAPAS 3D PARA NAVEGAÇÃO VIRTUAL: UMA ABORDAGEM COGNITIVA**. Tese—Curitiba: UNIVERSIDADE FEDERAL DO PARANÁ, 2012.

SCHWERING, A. Evaluation of a Semantic Similarity Measure for Natural Language Spatial Relations. In: WINTER, S. et al. (Eds.). **Spatial Information Theory**. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007. v. 4736p. 116–132.

SCHWERING, A.; RAUBAL, M. **Measuring Semantic Similarity Between Geospatial Conceptual Regions**. GeoSpatial Semantics. **Anais...** In: INTERNATIONAL CONFERENCE ON GEOSPATIAL SEMANTICS. 2005

SENA, V. A. S.; CLARO, D. B.; AMORIM, R.; LOPES, D. Similaridade Semântica na Composição de Sistemas de Informação através dos Serviços Web. **VI Simpósio Brasileiro de Sistemas de Informação (SBSI 2010)**, p. 12, 2010.

SENO, E. R. M.; NUNES, M. DAS G. V. N. Reconhecimento de Informações Comuns para a Fusão de Sentenças Comparáveis do Português. **Linguamática**, v. 1, p. 71–88, maio 2009.

SHARIF, A. R. B.; EGENHOFER, M. J.; MARK, D. M. Natural-Language Spatial Relations Between Linear and Areal Objects: The Topology and Metric of English-Language Terms. **International Journal of Geographical Information Science**, v. 12, n. 3, p. 215–245, 1998.

SHUSTERMAN, A.; LI, P. Frames of reference in spatial language acquisition. **Cognitive Psychology**, v. 88, p. 115–161, ago. 2016.

SORROWS, M. E.; HIRTLE, S. C. The Nature of Landmarks for Real and Electronic Spaces. **Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science**, Lecture Notes in Computer Science. v. 1661, p. 37–50, 1999.

SOUSA, M. R. F.; SILVA, E. L.; DIAS, G. A.; SILVA, M. A. T.; FREITAS, F. L. G.; AZEVEDO, R. R. InfoArch: uma ontologia para modelar o domínio da Arquitetura da Informação para Web. **Liinc em Revista - Laboratório Interdisciplinar em Informação e Conhecimento - XI Enancib: Ciência da Informação em foco**, v. 7, n. 1, p. 264–282, mar. 2011.

SOWA, J. F. **Semantic networks**. 1987.

STOCK, K. Determining Semantic Similarity of Behaviour Using Natural Semantic Metalanguage to Match User Objectives to Available Web Services. **Transactions in GIS**, v. 12, n. 6, p. 733–755, dez. 2008.

STOCK, K.; STOJANOVIC, T.; REITSMA, F.; OU, Y.; BISHR, J. O.; ROBERTSON, A. To ontologise or not to ontologise: An information model for a geospatial knowledge infrastructure. **Computers & Geosciences**, v. 45, p. 98–108, ago. 2012.

STOCK, K.; CIALONE, C.; PASLEY, R. C.; RANJIT, R.; POURABDOLLAH, A.; BROWN, M. Multilingual Natural Language Spatial Querying using Natural Semantic Metalanguage. 2013.

STOCK, K. **A Geometric Configuration Ontology to Support Spatial Querying**. In J. Huerta, S. Schade, & C. Granell (Eds.), Proceedings of AGILE 2014. **Anais... In: 17TH CONFERENCE ON GEOGRAPHIC INFORMATION SCIENCE**. 2014

STOCK, K. **Identifying Patterns in Geospatial Natural Language**. In: Poster at Conference on Spatial Information Theory. **Anais...2015**

STOCK, K.; HALL, M. The Role of Context in the Interpretation of Natural Language Location Descriptions. **Proceedings of Workshops and Posters at the 13th International Conference on Spatial Information Theory (COSIT 2017)**, Lecture Notes in Geoinformation and Cartography. p. 245–254, 2018.

STOCK, K. M.; ATKINSON, R.; HIGGINS, C.; SMALL, M.; WOOLF, A.; MILLARD, K.; ARCTUR, D. A semantic registry using a Feature Type Catalogue instead of ontologies to support spatial data infrastructures. **International Journal of Geographical Information Science**, v. 24, n. 2, p. 231–252, fev. 2010.

STOCK, K. M.; DELAZARI, L. S. Where am I?/Onde Estou? Automated Interpretation of Human Language Descriptions of Current Location. **Adfa**, p. 1, 2011.

STOCK, K.; PULLAR, D. Identifying Semantically Similar Elements in Heterogeneous Spatial Databases Using Predicate Logic Expressions. In: VČKOVSKI, A.; BRASSEL, K. E.; SCHEK, H.-J. (Eds.). **Interoperating Geographic Information Systems**. Lecture Notes in Computer Science. Berlin, Heidelberg: [s.n.]. v. 1580p. 231–252, 1999.

STOCK, K.; YOUSAF, J. Context-aware automated interpretation of elaborate natural language descriptions of location through learning from empirical data. **International Journal of Geographical Information Science**, v. 32, n. 6, p. 1087–1116, 3 jun. 2018.

SUCHAN, T. A.; BREWER, C. A. Qualitative Methods for Research on Mapmaking and Map Use. **Professional Geographer**, v. 52, n. 1, p. 145–154, 2000.

TALMY, L. HOW LANGUAGE STRUCTURES SPACE. **Spatial orientation**, n. Springer, p. 225–282, 1983.

TALMY, L. Lexicalization patterns: Semantic structure in lexical forms, in SHOPEN, INTRODUCTION 25 T. (1985, ed.), **Language typology and syntactic description**. Vol. III: Grammatical categories and the lexicon, Cambridge University Press, Cambridge: 57-149, 1985.

TALMY, L. **Toward a cognitive semantics**. [s.l.] MIT press, 2000. v. 2

TANASESCU, V. Spatial Semantics in Difference Spaces. In: WINTER, S. et al. (Eds.). **Spatial Information Theory**. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007. p. 96–115.

TEIXEIRA, J. De cá para lá e de aqui para aí: rede de valores semânticos dos marcadores espaciais cá/lá/(acolá) e aqui/aí/ali. In: RIO-TORTO, G. M. et al. (Eds.). **Estudos em homenagem ao Professor Doutor Mário Vilela**. Porto: Faculdade de Letras da Universidade do Porto, 2005. v. Ip. 449–460.

TEIXEIRA, J. Modelos Semânticos e Variação Diatópica. **I Encontro de Estudos Dialectológicos – Actas, Instituto Cultural de Ponta Delgada**, p. 363–380, 2006.

TEIXEIRA, J. SINONÍMIA E PROCESSOS DE IMPLICAÇÃO: ALGUMAS RELAÇÕES ENTRE ESPAÇO E TEMPO NO PORTUGUÊS EUROPEU. **Актуальные проблемы лузофонии. (Problemas atuais da lusofonia) Serta Lusitânica - Universidade de S. Petersburgo**, p. 56–77, 2013.

TENBRINK, T. **Space, time and the use of language: an investigation of relationships**. Berlin, Germany: [s.n.]. v. III, 2007.

TENBRINK, T. Discovering Spatiotemporal Concepts in Discourse. In: DANCYGIER, B. (Ed.). **The Cambridge Handbook of Cognitive Linguistics**. Cambridge: Cambridge University Press, 2017. p. 669–683.

TENBRINK, T.; BERGMANN, E.; KONIECZNY, L. Wayfinding and description strategies in an unfamiliar complex building. **Proceedings of the Annual Meeting of the Cognitive Science Society**, v. 33, p. 1262–1267, 2011.

TOMKO, M.; WINTER, S. RECONSTRUCTION OF SCENES FROM GEO-REFERENCED WEB RESOURCES. **Proceedings of SSC 2005 Spatial Intelligence, Innovation and Praxis: The national biennial Conference of the Spatial Science Institute**, p. 9, set. 2005.

TYLER, A. Spatial Language, Polysemy, and Cross-Linguistic Semantic Mismatches: Cognitive Linguistics Insights into Challenges for Second Language Learners. **Spatial Cognition & Computation**, v. 12, n. 4, p. 305–335, out. 2012.

TYLER, A.; EVANS, V. The Semantics of English Prepositions: Spatial scenes, cognition and the experiential basis of meaning. **Journal of Women s Health**, 2003.

URSINI, F. A. On the polysemy of Italian spatial prepositions. In: **Di tutti I colori. Studi linguistici**. Leiden University Repository: Roberta D'Alessandro Gabriele Iannàccaro Diana Passino Anna M. Thornton, 2017. p. 1–22.

VARANKA, D. E. Ontology Patterns for Complex Topographic Feature Types. **Cartography and Geographic Information Science**, v. 38, n. 2, p. 126–136, jan. 2011.

VARANKA, D.; JERRIS, T. J. COMPLEX TOPOGRAPHIC FEATURE ONTOLOGY PATTERNS. In: **Proceedings AutoCarto**, p. 5, 2010.

VASARDANI, M.; TIMPF, S.; WINTER, S.; TOMKO, M. **From Descriptions to Depictions: A Conceptual Framework**. Proceedings of COSIT 2013. **Anais...**: Lecture Notes in Computer Science. In: CONFERENCE ON SPATIAL INFORMATION THEORY. Heidelberg: Springer Verlag, 2013. Disponível em: <http://link.springer.com/10.1007/978-3-319-01790-7_17>. Acesso em: 4 mar. 2021

VIAENE, P.; VANCLOOSTER, A.; OOMS, K.; MAEYER, P. Thinking Aloud in Search of Landmark Characteristics in an Indoor Environment. In: **2014 Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS)**, IEEE. p. 103–110, 2014.

VINSON, N. G. **Design guidelines for landmarks to support navigation in virtual environments**. Proceedings of the SIGCHI conference on Human factors in computing systems the CHI is the limit - CHI '99. **Anais...** In: THE SIGCHI CONFERENCE. Pittsburgh, Pennsylvania, United States: ACM Press, 1999. Disponível em: <<http://portal.acm.org/citation.cfm?doid=302979.303062>>. Acesso em: 8 mar. 2021

WANG, X.; MATSAKIS, P.; TRICK, L.; NONNECKE, B.; VELTMAN, M. A Study on how Humans Describe Relative Positions of Image Objects. In: RUAS, A.; GOLD, C. (Eds.). **Headway in Spatial Data Handling**. Lecture Notes in Geoinformation and Cartography. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. p. 1–18.

WERNER, S.; KRIEG-BRUCKNER, B.; MALLOT, H. A.; SCHWEIZER, K.; FREKSA, C. Spatial Cognition: The Role of Landmark, Route, and Survey Knowledge in Human and Robot Navigation. **Informatik '97 Informatik als Innovationsmotor.**, n. Springer, Berlin, Heidelberg, p. 41–50, 1997.

WIERZBICKA, A. **Semantic Primitives**. Frankfurt: Athenäum, 1972.

WINTER, S.; HAMZEI, E.; WEGHE, N. V.; OOMS, K. **A Graph Representation for Verbal Indoor Route Descriptions**. Spatial Cognition XI. **Anais...** In: 11TH INTERNATIONAL CONFERENCE. Tübingen, Germany: Springer, Cham, 2018

YUE, P.; GONG, J.; DI, L.; HE, L.; WEI, Y. Integrating semantic web technologies and geospatial catalog services for geospatial information discovery and processing in cyberinfrastructure. **Geoinformatica**, v. 15, n. 2, p. 273–303, abr. 2011.

ZWARTS, J. Prepositional Aspect and the Algebra of Paths. **Linguistics and Philosophy**, v. 28, n. 6, p. 739–779, 1 dez. 2005.

ZWARTS, J. Spatial semantics: Modeling the meaning of prepositions. **Language and Linguistics Compass**, v. 11, n. 5, p. e12241, maio 2017.

ZWARTS, J.; WINTER, Y. Vector Space Semantics: A Model-Theoretic Analysis of Locative Prepositions. **Journal of Logic, Language and Information**, v. 9, n. 2, p. 169–211, 1 abr. 2000.

APÊNDICE 1 – TERMOS DE CONSENTIMENTO DOS TESTES

Teste *Outdoor*

FORMULÁRIO DO TERMO DE COMPROMISSO

Proponentes:

Profª. Drª. Luciene Stamato Delazari – UFPR.

Prof. Dr. Marcio Schmidt – UFU.

Profª. Drª. Kristin Stock – Massey University

Prezado Participante,

Neste termo são apresentadas as garantias e condições que serão dadas a você, caso aceite participar desta pesquisa. As análises das informações contidas nas respostas dos questionários comporão os resultados da pesquisa, cujo objetivo é avaliar como as pessoas descrevem o ambiente.

Desta forma, nós garantimos que **(1)** seus dados pessoais, bem como elementos que permitam sua identificação por terceiros, serão mantidos em sigilo; **(2)** você não responderá perguntas de cunho pessoal; **(3)** você pode interromper os testes.

Ao assinar este termo você concorda com **(1)** a utilização dos dados presentes nos questionários com finalidade de análise para pesquisa científica; **(2)** está ciente das condições para a realização dos testes; **(3)** está ciente das garantias a você dadas.

Caso não tenha compreendido qualquer item, não hesite em perguntar.

Desde já agradecemos sua disposição.

Eu li e concordo com os termos descritos



Teste Indoor

Seção 1 de 3

Estudo das Relações Espaciais existentes no Ambiente Indoor

TERMO DE COMPROMISSO

Proponentes:

Ms^a. Cristiane K. M. Fagundes – Doutoranda, Programa de Pós-Graduação em Ciências Geodésicas, UFPR

Ms^a. Rhaíssa V. Sarot – Doutoranda, Programa de Pós-Graduação em Ciências Geodésicas, UFPR

Prof^a. Dr^a. Luciene Stamato Delazari – Docente, Programa de Pós-Graduação em Ciências Geodésicas, UFPR

Prezado Participante,

Neste termo são apresentadas as garantias e condições que serão dadas a você, caso aceite participar desta pesquisa. As análises das informações contidas nas respostas dos questionários comporão os resultados da pesquisa, cujo objetivo é a análise das relações espaciais pertencentes aos ambientes indoor. Pretende-se minimizar os problemas relacionados à orientação e navegação espacial do usuário em se locomover nestes ambientes.

Desta forma, nós garantimos que :

- (1) seus dados pessoais, bem como elementos que permitam sua identificação por terceiros, serão mantidos em sigilo;
- (2) você não responderá perguntas de cunho pessoal;
- (3) você pode interromper os testes (tempo médio de duração: 35 min).

Ao assinar este termo você concorda com (1) a utilização dos dados presentes nos questionários com finalidade de análise para pesquisa científica; (2) está ciente das condições para a realização dos testes; (3) está ciente das garantias a você dadas.

Caso não tenha compreendido qualquer item, não hesite em perguntar.

Desde já agradecemos sua disposição.

Endereço de e-mail *

Endereço de e-mail válido

Este formulário coleta endereços de e-mail. [Alterar configurações](#)

Eu li e concordo com os termos descritos *

Sim

Após a seção 1 Continuar para a próxima seção

Teste Similaridade

Pesquisa UFPR: Relações Espaciais

Proponentes:

Cristiane Kutianski Marchi Fagundes – doutoranda – e-mail: cristiane.fagundes@ufpr.br

Prof^a. Dr^a. Luciene Stamato Delazari – docente – e-mail: luciene@ufpr.br

Programa de Pós-Graduação em Ciências Geodésicas da UFPR.

*Obrigatório

Termo de Consentimento: Prezado (a), a pesquisa abaixo é parte de um estudo que tem por objetivo analisar as relações espaciais de descrições em linguagem natural e suas similaridades semânticas. Para a realização da pesquisa, as proponentes garantem que: o usuário não terá nenhum custo, nenhum dado pessoal será divulgado e somente os pesquisadores envolvidos terão acesso aos dados coletados. A participação é voluntária e anônima. A recusa não acarretará em qualquer penalidade. Caso sinta algum desconforto, poderá recusar ou interromper a participação a qualquer momento. Ao aceitar este termo, você concorda em participar da pesquisa. Reiteramos que a utilização dos dados coletados será exclusivamente para fins acadêmicos. *

Aceito

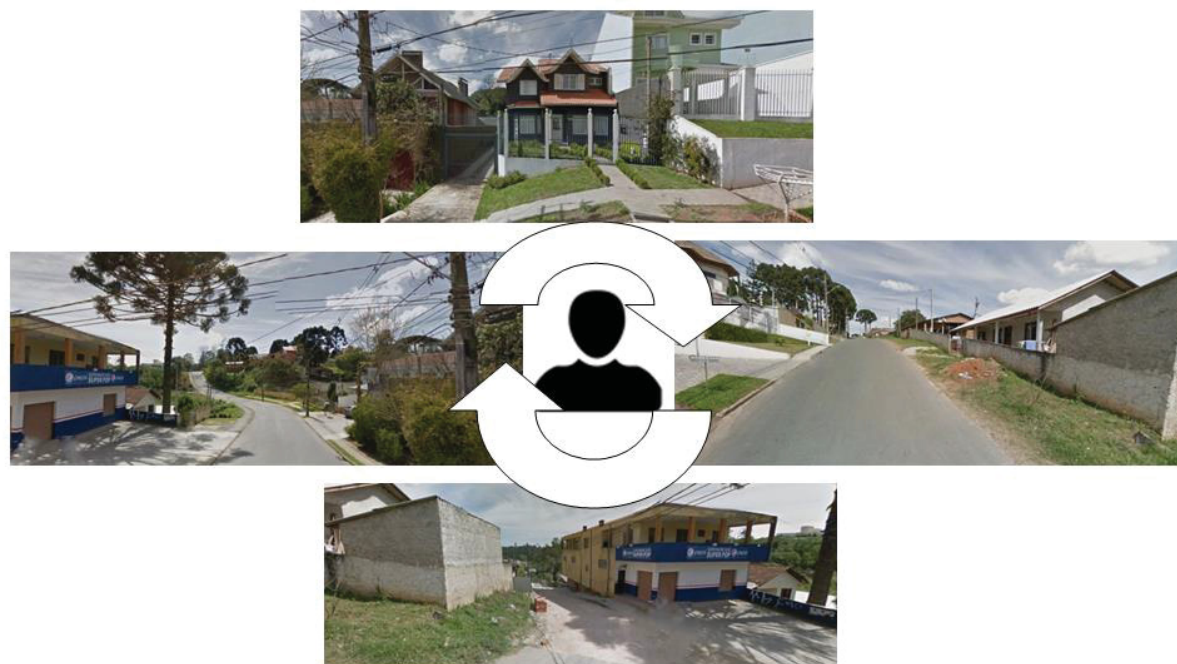
Não aceito

APÊNDICE 2 – FOTOGRAFIAS DAS LOCALIDADES DO TESTE *OUTDOOR*

BRA1: Praça Santos Dumond, bairro: Centro - Curitiba-PR



BRA2: Rua Eugênio Flôr, 1091, bairro: Abranches - Curitiba-PR



BRA3: Praça Tubal Vilela, bairro: Centro - Uberlândia-MG



BRA4: Felisberto Carrijo, 1059, bairro: Centro – Uberlândia-MG



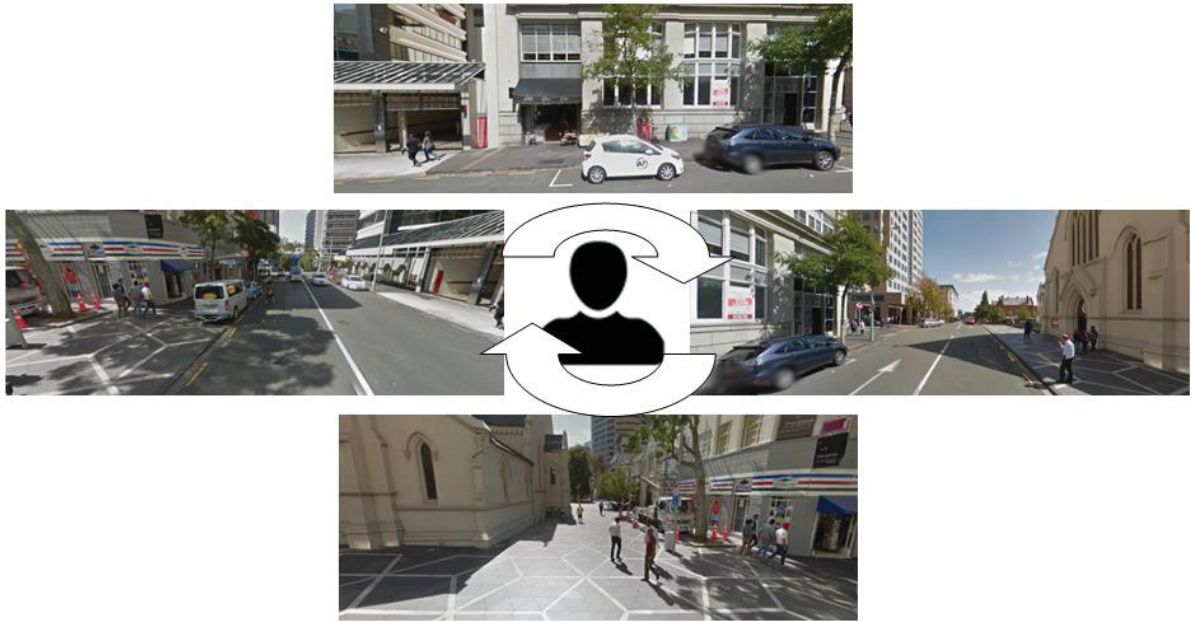
NZ1: 309, Hibiscus Coast Hwy, Orewa – Auckland-NZ



NZ2: State Highway 1, Puhoi – Auckland-NZ



NZ3: 37, Wyndham St, Downtown – Auckland-NZ



NZ4: 79, Wainui Ave – Auckland-NZ



NZ5: 147, Newton RD, Eden Terrace – Auckland-NZ

