


RESSALVA

Atendendo solicitação do(a) autor(a), o texto completo desta dissertação será disponibilizado somente a partir de 06/12/2023.

unesp  **UNIVERSIDADE ESTADUAL PAULISTA**
“JÚLIO DE MESQUITA FILHO”
Faculdade de Ciências e Letras
Campus de Araraquara - SP

MILENA APARECIDA DE ALMEIDA

A SOCIOLINGUÍSTICA DAS ESTRATÉGIAS DE RELATIVIZAÇÃO EM CONTEXTOS LOCATIVOS NA FALA PAULISTA



ARARAQUARA – S.P.
2022

MILENA APARECIDA DE ALMEIDA

A SOCIOLINGUÍSTICA DAS ESTRATÉGIAS DE RELATIVIZAÇÃO EM CONTEXTOS LOCATIVOS NA FALA PAULISTA

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Linguística e Língua Portuguesa da Faculdade de Ciências e Letras – Unesp/Araraquara, como requisito para obtenção do título de Mestre em Linguística e Língua Portuguesa.

Linha de pesquisa: Análise Fonológica, Morfossintática, Semântica e Pragmática

Orientadora: Prof.^a Dr.^a Rosane de Andrade Berlinck

Co-orientador: Prof. PhD Stephen Levey

Bolsa: FAPESP (Processo nº 2020/00593-5)

ARARAQUARA – S.P.
2022

A447s Almeida, Milena Aparecida de
A sociolinguística das estratégias de relativização em contextos locativos na fala paulista / Milena Aparecida de Almeida. -- Araraquara, 2022
132 f. : tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências e Letras, Araraquara
Orientadora: Rosane de Andrade Berlinck
Coorientador: Stephen Levey

1. Sociolinguística. 2. Sintaxe. 3. Estratégias de relativização. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências e Letras, Araraquara. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

MILENA APARECIDA DE ALMEIDA

A SOCIOLINGUÍSTICA DAS ESTRATÉGIAS DE RELATIVIZAÇÃO EM CONTEXTOS LOCATIVOS NA FALA PAULISTA

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Linguística e Língua Portuguesa da Faculdade de Ciências e Letras – Unesp/Araraquara, como requisito para obtenção do título de Mestre em Linguística e Língua Portuguesa.

Linha de pesquisa: Análise Fonológica, Morfossintática, Semântica e Pragmática

Orientadora: Prof.^a Dr.^a Rosane de Andrade Berlinck

Co-orientador: Prof. PhD Stephen Levey

Bolsa: FAPESP (Processo nº 2020/00593-5)

Data da defesa: 11/02/2022

MEMBROS COMPONENTES DA BANCA EXAMINADORA:

Presidente e Orientador: Prof.^a Dr.^a Rosane de Andrade Berlinck

Universidade Estadual Paulista – UNESP / FCLAR

Membro Titular:

Prof.^a Dr.^a Livia Oushiro

Universidade de Campinas – UNICAMP/ IEL

Membro Titular:

Prof.^a Dr.^a Edair Görski

Universidade Federal de Santa Catarina - UFSC

Local: Universidade Estadual Paulista
Faculdade de Ciências e Letras
UNESP – Campus de Araraquara

Àqueles que me ensinaram sobre afeto incondicional e disciplina,
Marlene e Vandecir

AGRADECIMENTOS

Dedicar essa dissertação a minha família parece simplório mediante todo o esforço mobilizado para que esse momento fosse possível. Dessa forma, tento expressar toda a minha gratidão nessas linhas aos meus pais, Marlene e Vandecir, minhas inspirações diárias nessa batalha, por nunca duvidarem do meu potencial e acreditarem fielmente que a filha de um casal “da roça”, de uma cidade pequena, com poucas chances de alfabetização, pudesse um dia receber um título de Mestre. Não menos importante, agradeço a minha irmã, Andresa, por me compreender e me apoiar em tantos momentos difíceis.

À minha orientadora, Rosane, a quem constantemente estou testando a paciência com inúmeros projetos e entregas em cima do prazo, eu sou eternamente grata. Nosso caminho se cruzou em 2016, quando era apenas uma caloura ansiosa demais, e desde então você me deu todas as ferramentas, incentivos e liberdade para que eu trilhasse meu próprio caminho dentro da academia. Obrigada pela confiança no meu trabalho, pelos conselhos constantes e por saber dosar minha expectativa com a realidade.

Ao meu coorientador, Stephen Levey, um dos pesquisadores mais atenciosos que encontrei em toda a minha trajetória, agradeço por todo o apoio e encorajamento nesses dois anos.

Minha jornada acadêmica, até o presente momento, não teria sido tão incrível e divertida se não fosse por essas duas figuras: Marcus e Letícia. Ao meu mentor, guia espiritual e companheiro de vida, Marcus, sou extremamente grata. Nunca conseguirei colocar em palavras como eu sou agradecida ao universo pela nossa amizade, por todo o tempo que você demanda apenas para entender minha pesquisa, me auxiliar das mais diversas formas possíveis, nem que para isso seja necessário deixar de lado as redações (e o dinheiro) para assistir vídeos de *dragqueen* comigo. Eu te amo e quero reconhecer publicamente que metade dessa dissertação não existiria sem a sua brilhante mente, tão desesperadoramente incrível em produzir um diálogo com uma perspectiva que você não domina! A sua radiação faz meus dias muito mais felizes, mesmo queimando meus bolos de consolação.

A minha parceira de dificuldades, Letícia, eu gostaria de agradecer e te entregar o mundo! Uma pessoa extremamente perturbada e ansiosa, que de alguma forma se assemelha muito a minha própria perturbação e ansiedade. Dividir essa jornada, dia após dia, com você, é o que torna esse caminho mais leve. Inclusive dentro de uma pandemia, não podendo de ver e ouvir pessoalmente desde o dia que nos intitulamos mestrandas. Temos muitas comemorações, tristezas e inseguranças para compartilhar, e sei que nosso caminho ainda é longo. Obrigada por tudo, miga!

Sou extremamente grata ao meu fã número 1, Raphael, a pessoa mais incrível e especial que encontrei no meio de um caminho muito turbulento. Muito obrigada por ser tão paciente, por querer desesperadamente entender meu cotidiano tão diferente do seu, por me ensinar a dosar minha ansiedade, expectativa e prioridades, e por me apoiar incondicionalmente em qualquer decisão profissional, mesmo que isso impacte seu próprio dia a dia. Essa dissertação tem um pedacinho seu, assim como você tem um pedacinho do meu coração. Te amo!

Agradeço aos meus parceiros roraimenses, Eliabe e Fabrício, por compartilharem tanto conhecimento, por acreditarem no meu potencial e por me animarem nos momentos difíceis

com fofocas e histórias absurdas. Apesar de não dividirmos açaí com tanta constância, meu amor e carinho por vocês não diminuiu, apenas multiplicou. Obrigada pela paciência!

À minha banca de qualificação e defesa, Livia Oushiro, Roberto Camacho e Edair Görski, meu muito obrigada por todas as colocações que tornaram meu trabalho muito mais coeso! Em especial, toda a minha gratidão à Livia Oushiro, que tem me assistido e acompanhado desde a graduação e está sempre disponível para as dúvidas sobre o R.

Aos meus amigos do Núcleo de Pesquisa em Sociolinguística de Araraquara, agradeço imensamente por toda a escuta carinhosa e as críticas produtivas que compartilhamos no decorrer dos últimos anos!

Por fim, agradeço à FAPESP (Processo nº 2020/00593-5) e à CAPES pelo apoio financeiro, fundamental para o desenvolvimento desta pesquisa.

“Vamos atentar para a sintaxe dos paulistas...”

(Língua – Caetano Veloso)

RESUMO

No contexto dos estudos sobre relativização em português, há poucos estudos específicos sobre o contexto locativo, que investiguem a variação entre as estratégias possíveis. Por outro lado, estudos recentes identificaram um uso não locativo de **onde**, apontando para um processo de gramaticalização do item na fala de diferentes regiões (SOUZA, 2003; BRAGA, MANFILI, 2004; ALMEIDA, 2020). Tendo em conta esse cenário, o objetivo do presente estudo é investigar o processo de variação linguística de construções relativas no domínio de origem do pronome relativo **onde**, isto é, na expressão do valor locativo na fala paulista. Essa caracterização é uma primeira etapa de investigação, que nos levará, em estudo posterior, a avaliar a possível correlação entre a variação no domínio de origem e os processos de multifuncionalidade de **onde**. Temos como envelope de variação as estratégias de relativização do Português Brasileiro (AMARAL, 1920; MOLLICA, 1977; TARALLO, 1983; CORRÊA, 1998; BISPO, 2003; 2009) – isto é, **padrão (com e sem preposição)**, **copiadora** e **cortadora**. Adotamos, assim, a abordagem teórico-metodológica da Teoria de Variação e Mudança Linguística (WEINREICH, LABOV, HERZOG, 2006 [1968]; LABOV, 1972) para análise do *corpus* construído tendo por base os bancos de dados IBORUNA, do projeto ALIP (Gonçalves, s.d) e Projeto SP2010: Amostra da fala paulistana (MENDES, 2013) contabilizando 1007 relativas. Estas estão distribuídas em: 3 dados de **padrão com preposição**, 35 dados de **copiadora**, 249 dados de **cortadora** e 267 dados de **padrão com onde** provenientes do banco de dados IBORUNA; e 10 dados de **padrão com preposição**, 15 dados de **copiadora**, 128 dados de **padrão com onde** e 300 dados de **cortadora** provenientes do banco de dados Projeto SP2010. Por conta da baixa frequência em nosso *corpus*, as estratégias **padrão com preposição** e **copiadora** receberam um tratamento qualitativo, enquanto uma análise quantitativa foi produzida para compreender o comportamento das estratégias **padrão com onde** e **cortadora**. Os fatores linguísticos que norteiam a pesquisa são: função sintática do pronome relativo; distância entre o SN e a oração relativa; classificação da relativa; identidade e tipo do locativo; definitude, especificidade e status informacional do SN antecedente. Os fatores extralinguísticos controlados são: sexo/gênero, escolaridade e faixa etária do falante (presentes nos dois bancos); tipologia textual (presente no IBORUNA); região e zona de residência do falante em São Paulo e roteiro da entrevista (presentes no SP2010). Em nossa análise, averiguamos, por meio de uma regressão logística de efeitos mistos, que há um desfavorecimento do uso da **padrão com onde** para falantes da primeira faixa etária, com níveis de escolaridade mais baixos e quando o antecedente se refere a um local, para o banco de dados IBORUNA, e um desfavorecimento do uso da **padrão com onde** para falantes da segunda e terceira faixa etária com ensino médio, moradores da região centro velho e em contextos indefinidos, para o banco de dados Projeto SP2010. Concluímos que a descrição desse contexto específico de relativização revela padrões diferentes em relação às tendências gerais (de predomínio de estratégias não padrão), já que a **padrão com onde** está ainda bem presente. Além disso, as variáveis sociais se mostraram muito significativas na explicação da variação, algo que contribui com uma discussão dentro do modelo teórico sobre o papel dessas variáveis em fenômenos morfológicos e sintáticos, somando-se a alguns estudos que mostram que pode, sim, haver influência desses aspectos em níveis mais altos. Ou seja, que é possível falar em variáveis sociolinguísticas (propriamente ditas) nesses níveis. Por fim, considerando que a relativização é um mecanismo presente em muitas línguas, para o qual se busca identificar princípios gerais, o estudo pode contribuir para essa discussão em âmbito interlinguístico.

Palavras – chave: pronome relativo onde; estratégias de relativização; português paulista; variável sintática.

ABSTRACT

In the context of studies on relativization in Portuguese, there are few specific studies on the locative context that investigate the variation between possible strategies. On the other hand, recent studies have identified a non-locative use of **onde**, pointing to a process of grammaticalization of the item in the speech of different regions (SOUZA, 2003; BRAGA, MANFILI, 2004; ALMEIDA, 2020). Taking this scenario into account, the objective of the present study is to investigate the process of linguistic variation of relative constructions in the domain of origin of the relative pronoun **onde**, that is, in the expression of the locative value in São Paulo speech. This characterization is a first stage of investigation, which will lead us, in a later study, to evaluate the possible correlation between the variation in the domain of origin and the multifunctionality processes from **onde**. Our envelope of variation includes the relativization strategies used in Brazilian Portuguese (AMARAL, 1920; MOLLICA, 1977; TARALLO, 1983; CORRÊA, 1998; BISPO, 2003; 2009) – that is, **standard (with or without preposition)**, **resumptive pronoun** and **pp-chopping**. We adopted the theoretical-methodological approach of the Theory of Language Variation and Change (WEINREICH, LABOV, HERZOG, 2006 [1968]; LABOV, 1972) to analyze the corpus built from the IBORUNA database (ALIP project (Gonçalves, n.d.)) and from Projeto SP2010: Amostra da fala paulistana (MENDES, 2013). The corpus counts 1007 tokens of relatives. These are distributed as: 3 occurrences of **standard with preposition**, 35 occurrences of **resumptive pronoun**, 249 occurrences of **pp-chopping** and 267 occurrences **standard with onde** from the IBORUNA database; and 10 occurrences of **standard with preposition**, 15 occurrences of **resumptive pronoun**, 128 occurrences of **standard with onde** and 300 data of **pp-chopping** from the SP2010 database. Due to the low frequency in our corpus, the **standard with preposition** and **resumptive pronoun** strategies received a qualitative treatment, while a quantitative analysis is being produced to understand the behavior of the **standard with onde** and **pp-chopping** strategies. The linguistic factors that guide the research are: syntactic function of the relative pronoun; distance between the NP and the relative clause; relative classification; identity and type of locative; definiteness, specificity and informational status of the antecedent NP. The controlled extralinguistic factors are: sex/gender, education and age group of the speaker (present in both database); textual type (present in IBORUNA); speaker's region and area of residence in São Paulo and interview script (present in SP2010). Our statistical analyses verified, through a mixed effects logistic regression, that the use of **standard with onde** is disfavored by speakers of the first age group, with lower levels of education and when the antecedent refers to a location, for the IBORUNA database, and the use of **standard with onde** is disfavored by speakers of the second and third age groups with high school, downtown residents and in undefined contexts, for SP2010. We conclude that the description of this specific context of relativization reveals different patterns in relation to the general trends (prevalence of non-standard strategies), since the **pattern with onde** is still very present. In addition, social variables proved to be very significant in explaining variation, something that contributes to a discussion within the theoretical model about the role of these variables in morphological and syntactic phenomena, in addition to some studies that show that there may, indeed, be influence of these aspects at higher levels. In other words, it is possible to speak of sociolinguistic variables (properly so-called) at these levels. Finally, considering that relativization is a mechanism present in many languages, for which it seeks to identify general principles, the study can contribute to this discussion in an interlinguistic scope

Keywords: relative pronoun **onde**; relativization strategies; Paulista Portuguese; syntactic variable.

LISTA DE FIGURAS

Figura 1	Fórmula da subordinação adjetiva no Português Brasileiro	29
Figura 2	A taxonomia dos valores da Familiaridade Assumida	64
Figura 3	Divisão da cidade de São Paulo entre regiões mais centrais e mais periféricas	72
Figura 4	Proporção de uso das estratégias de relativização no banco de dados IBORUNA	75
Figura 5	Proporção de uso das estratégias de relativização no banco de dados Projeto SP2010	76
Figura 6	Nuvem de palavras da identidade lexical do verbo na oração relativa no banco de dados IBORUNA e Projeto SP2010, respectivamente	78
Figura 7	Proporção de uso das estratégias de relativização segundo a faixa etária do falante no banco de dados IBORUNA	81
Figura 8	Proporção de uso das estratégias de relativização segundo a faixa etária do falante no banco de dados SP2010	82
Figura 9	Proporção e frequência das estratégias de relativização segundo a escolaridade do falante no banco de dados IBORUNA	84
Figura 10	Proporção e frequência das estratégias de relativização segundo a escolaridade do falante no banco de dados SP2010	84
Figura 11	Árvore de classificação condicional para o uso das estratégias de relativização segundo a faixa etária e escolaridade do falante no banco de dados IBORUNA	88
Figura 12	Árvore de classificação condicional para o uso das estratégias de relativização segundo a faixa etária e escolaridade do falante no banco de dados SP2010	88
Figura 13	Proporção de uso das estratégias de relativização segundo a região de residência do falante	90
Figura 14	Proporção de uso das estratégias de relativização segundo a zona de residência do falante no banco de dados SP2010	91
Figura 15	Proporção de uso das estratégias de relativização segundo a região de residência e a zona de residência do falante no banco de dados SP2010	92

Figura 16	Proporção de uso das estratégias de relativização segundo a identidade semântica do antecedente no banco de dados IBORUNA	93
Figura 17	Proporção de uso das estratégias de relativização segundo a identidade semântica do antecedente no banco de dados SP2010	94
Figura 18	Proporção de uso das estratégias de relativização segundo a definitude do antecedente no banco de dados Projeto SP2010	95
Figura 19	Proporção de uso das estratégias de relativização segundo a definitude e a especificidade do antecedente no banco de dados SP2010	96
Figura 20	Proporção de uso das estratégias de relativização segundo a definitude e o status informacional do antecedente no banco de dados SP2010	98
Figura 21	Proporção de uso das estratégias de relativização segundo a definitude do antecedente no banco de dados IBORUNA	99
Figura 22	Proporção de uso das estratégias de relativização segundo a definitude e a especificidade do antecedente no banco de dados IBORUNA	100
Figura 23	Proporção de uso das estratégias de relativização segundo o status informacional do antecedente no banco de dados IBORUNA	102
Figura 24	Proporção de uso das estratégias de relativização segundo a adjacência entre o sintagma nominal antecedente e a oração relativa no banco de dados IBORUNA	104
Figura 25	Proporção de uso das estratégias de relativização segundo a adjacência entre o sintagma nominal antecedente e a oração relativa no banco de dados SP2010	105
Figura 26	Frequência e proporção de uso das estratégias de relativização segundo a classificação semântica da relativa no banco de dados IBORUNA	106
Figura 27	Frequência e proporção de uso das estratégias de relativização segundo a classificação semântica da relativa no banco de dados SP2010	106
Figura 28	Frequência e proporção de uso das estratégias de relativização segundo classe morfológica do locativo no banco de dados IBORUNA	108
Figura 29	Frequência e proporção de uso das estratégias de relativização segundo classe morfológica do locativo no banco de dados SP2010	108

Figura 30	Proporção de uso das estratégias de relativização segundo a função sintática do pronome relativo no banco de dados IBORUNA	109
Figura 31	Proporção de uso das estratégias de relativização segundo a função sintática do pronome relativo no banco de dados SP2010	110
Figura 32	Frequência e proporção de uso das estratégias de relativização segundo o sexo do falante no banco de dados IBORUNA	112
Figura 33	Frequência e proporção de uso das estratégias de relativização segundo o sexo do falante no banco de dados SP2010	112
Figura 34	Frequência e proporção de uso das estratégias de relativização segundo o tipo textual no banco de dados IBORUNA	113
Figura 35	Frequência e proporção de uso das estratégias de relativização segundo o roteiro de entrevista no banco de dados Projeto SP2010	114
Figura 36	<i>Continuum</i> de marcação (givoniana) aplicado às estratégias de relativização	116
Figura 37	Proporção de uso das estratégias de relativização segundo a faixa etária e a escolaridade do falante no banco de dados IBORUNA	119
Figura 38	Proporção de uso das estratégias de relativização segundo a faixa etária e a escolaridade do falante no banco de dados SP2010	120

LISTA DE TABELAS

Tabela 1	Descrição do envelope de variação	20
Tabela 2	Frequência e proporção de uso de onde “não locativo” segundo escolaridade no bando de dados IBORUNA	26
Tabela 3	Frequência e proporção de uso de onde “não locativo” segundo a faixa etária do falante no bando de dados IBORUNA	27
Tabela 4	Distribuição da proporção de uso das estratégias de relativização de acordo com estratégia em 5 funções sintáticas	36
Tabela 5	Porcentagem de uso das estratégias copiadora, cortadora e <i> piedpiping </i> nas três posições sintáticas mais baixas para as três classes sociais	36
Tabela 6	Distribuição da proporção de uso das estratégias de relativização <i> piedpiping </i> , copiadora e cortadora através do tempo	37
Tabela 7	Relativas no corpus <i> D&G Rio de Janeiro </i> e <i> D&G Natal </i> por estratégia e nível de escolaridade	39
Tabela 8	Distribuição das estratégias de relativização em ambientes preposicionados em Barros (2000)	41
Tabela 9	Distribuição das estratégias de relativização segundo a função sintática do pronome relativo em Burgos (2003)	42
Tabela 10	<i> Coding protocol </i> utilizado para as variáveis linguísticas	53
Tabela 11	Padrões utilizados na análise da definitude do antecedente	62
Tabela 12	<i> Coding protocol </i> utilizado para as variáveis extralinguísticas	68
Tabela 13	Análise de regressão de efeitos mistos da estratégia de relativização padrão com onde no banco de dados IBORUNA	79
Tabela 14	Análise de regressão de efeitos mistos da estratégia de relativização padrão com onde no banco de dados Projeto SP2010	80
Tabela 15	Frequência e proporção de uso das estratégias de relativização segundo a adjacência entre o sintagma nominal antecedente e a oração relativa no banco de dados IBORUNA	116
Tabela 16	Frequência e proporção de uso das estratégias de relativização segundo o a adjacência entre o sintagma nominal antecedente e a oração relativa no banco de dados SP2010	117

Tabela 17	Frequência e proporção de uso das estratégias de relativização segundo a faixa etária do falante no banco de dados IBORUNA	118
Tabela 18	Frequência e proporção de uso das estratégias de relativização segundo a faixa etária do falante no banco de dados SP2010	118

SUMÁRIO

1 INTRODUÇÃO	17
2 FUNDAMENTAÇÃO TEÓRICA	22
2.1 POR <i>ONDE</i> ANDA VOCÊ? UMA BREVE CONTEXTUALIZAÇÃO	22
2.1.1 Uma perspectiva interlinguística	22
2.1.2 Uma investigação paulista	24
2.2 O OBJETO DE ESTUDO	28
2.2.1 Tradição normativa	30
2.2.2 Tradição descritiva	32
2.2.3 As estratégias de relativização no Português	33
2.3 PRESSUPOSTOS TEÓRICO-METODOLÓGICOS	42
2.3.1 A Teoria de Variação e Mudança Linguística	42
2.3.1.1 A problemática da variável sintática	45
3 DECISÕES TEÓRICO-METODOLÓGICAS	48
3.1 VARIÁVEL DEPENDENTE	50
3.2 VARIÁVEIS LINGUÍSTICAS	51
3.2.1 Função sintática do pronome relativo	53
3.2.2 Adjacência entre o sintagma nominal antecedente e a oração relativa	55
3.2.3 Classe morfológica do locativo	57
3.2.4 Identidade semântica (ou tipo) do locativo	57
3.2.5 Classificação semântica da relativa	58
3.2.6 Definitude, especificidade e status informacional do antecedente	59
3.3 VARIÁVEIS EXTRALINGUÍSTICAS	66
3.3.1 Sexo/gênero	67
3.3.2 Escolaridade	68
3.3.3 Faixa etária	69
3.3.4 Região e zona de residência de São Paulo	70
3.3.5 Tipologia textual	72
3.3.6 Roteiro da entrevista	72

4 A SINTAXE DOS PAULISTAS: ANÁLISE DE PRODUÇÃO	73
4.1 DISTRIBUIÇÃO DE DADOS DAS VARIÁVEIS CORRELACIONADAS	75
4.1.1 Síntese	100
4.2 DISTRIBUIÇÃO DE DADOS DAS VARIÁVEIS NÃO CORRELAIONADAS	102
4.2.1 Adjacência entre o sintagma nominal antecedente e a oração relativa	102
4.2.2 Classificação semântica da relativa	104
4.2.3 Classe morfológica do locativo	106
4.2.4 Função sintática do pronome relativo	107
4.2.5 Sexo/gênero do falante	110
4.2.6 Tipologia Textual	112
4.2.7 Roteiro de entrevista	113
4.3 O CASO DA COPIADORA E PADRÃO COM PREPOSIÇÃO	114
5 CONSIDERAÇÕES FINAIS	121
REFERÊNCIAS	125

1 INTRODUÇÃO

Lehmann (2002), a partir de uma abordagem funcionalista da linguagem, retoma uma proposta do âmbito dos Estudos do Léxico, utilizada desde o século XIX, com o objetivo de estender tais noções para outros níveis linguísticos de análise. Assim, as formas linguísticas – em processo de competição ou não – podem ser investigadas a partir de duas perspectivas: a semasiológica e a onomasiológica. A primeira tem por base uma linha de estudos que parte de uma determinada forma (seja ela uma palavra, uma expressão ou uma construção) para o mapeamento de seus sentidos (ou funções); enquanto a segunda parte de um conceito (ou função) para identificar que formas o expressam na língua subjacente às construções (ou formas) investigadas, descrevendo como é expresso na língua alvo.

Essa distinção se mostra muito produtiva para o estudo de fenômenos de variação em níveis gramaticais mais altos dentro da Teoria de Variação e Mudança Linguística (WEINREICH, LABOV, HERZOG, 2006 [1968]; LABOV, 1972, 1982, 1994, 2001, 2010). Isso, pois, apesar de assumir que a variação é constitutiva da língua, podendo afetar qualquer nível linguístico, há uma problemática à teoria no que diz respeito aos fenômenos linguísticos acima do nível fonético-fonológico (LAVANDERA, 1978; LABOV, 1978; MILROY, GORDON, 2003; FREITAG, 2009). Diante dessa questão, a proposta de análise de Lehmann (2002) contribui para sanar um dos problemas encontrados¹, norteados os caminhos que o pesquisador pode percorrer para investigar uma ou mais formas linguísticas de nível gramatical mais alto.

Um exemplo de trabalho morfossintático que toma essa distinção é a investigação da expressão do *antepresente* e do *passado absoluto* no espanhol de Madri, Buenos Aires e San Miguel de Tucumán, de Araújo (2017). O autor buscou identificar como esses dois valores se materializam, de forma variável, por meio das formas verbais do *pretérito perfecto compuesto* (como em *he escrito*) e do *pretérito perfecto simple* (como em *escribí*), o que caracteriza uma perspectiva onomasiológica. Para chegar a essa análise, Araújo se vale da descrição semasiológica que foi feita anteriormente sobre os dois tempos verbais (como, por exemplo, GILI GAYA (1970); RAE (1986) e TORREGO (2002)), tendo em vista uma melhor compreensão do fenômeno em questão.

Entre várias possíveis aplicações, a conjugação das duas perspectivas indicadas por Lehmann (2002) se revela muito produtiva para o estudo da forma e função do pronome relativo

¹ A discussão completa se encontra na seção 2.3.1.1: *A problemática da variável sintática*

onde no Português Brasileiro, nosso objeto de estudo. O modo como os falantes têm-se utilizado do pronome para cumprir determinadas funções comunicativas que fogem da prescrição gramatical – isto é, o valor locativo – tem despertado o interesse de vários pesquisadores (BRAGA, MANFILI, 2004; LIMA, COROA, 2013; LIMA, 2007; SOUZA, 2003; ALMEIDA, BERLINCK, 2019; ALMEIDA, 2020). Alguns desses usos podem ser observados nos seguintes exemplos:

- (1.1) “terríveis... terrível **os últimos dois anos onde eu passei a desconfiá(r) que ele tinha...** caso com uma colega de trabalho” (AC-101; L. 58-61)
- (1.2) “o povo não tem votado direito... e... o país os municípios os estados... não têm sido bem sucedido... em algumas eleições... vide:: **a eleição do... Fernando Collor... onde ele pregô(u) tanto...** e depois foi... deu no que deu... ele o povo brasile(i)ro naquela... esperança da salvação que o povo vive até hoje...” (AC-113; RO: L. 218-222)

Em (1.1) o pronome relativo **onde** é usado na construção para uma referência ao sintagma “os últimos dois anos”, ou seja, a referência por ele efetuada é temporal, e não estritamente espacial. Em (1.2) o pronome adquire um valor diferente, já que, ao efetuar o processo de referência ao sintagma anterior, ele não produz o efeito locativo prescrito ao introduzir um evento, isto é, referente “a eleição do Fernando Collor”.

Adotando uma perspectiva semasiológica² (ALMEIDA, 2019), investigamos as funções que o pronome tem exercido na língua, considerando-as como parte de um possível processo de gramaticalização, a partir de um *corpus* composto exclusivamente da fala paulista de duas comunidades – uma do interior de São Paulo e outra da capital. Os usos não locativos de **onde** encontrados, tal qual (1.1) e (1.2), corroboram a literatura sobre o fenômeno e comprovam sua multifuncionalidade. Em especial, é necessário salientar que as ocorrências mais frequentes foram, justamente, da categoria semântico-cognitiva mais gramaticalizada, denominada por nós como “texto”.

- (1.3) às vez a mãe bebe demais o pai bebe demais e vai querê(r) í(r) falá(r) po filho – “num pode” –?... tipo uma que ele vai falá(r) – “por que que você pode e eu não posso?” – aí vai começá(r) a fazê(r) escondido... **e é onde o mundo tá do jeito que TÁ...** (AC-024; RO: L. 418-420)

Em (1.3), o falante parte de uma dada situação, descrevendo-a para que seu interlocutor compreenda a argumentação e conclui introduzindo uma oração encabeçada com **onde**. Não há

² Descrita em mais detalhes na seção 2: *Por onde anda você? Uma breve contextualização.*

um antecedente explícito, um núcleo ao qual uma oração relativa se liga; a articulação se dá de forma mais frouxa, com função semelhante ao de uma conjunção em construções adverbiais, expressando relações de causa, condição e tempo. O que observamos nesse tipo de ocorrência é um uso de **onde** que apresenta um caráter sintático e semântico diferente daquele considerado normativo, uma vez que não se configura mais como um pronome relativo, aproximando-se mais de um marcador discursivo, ou uma conjunção, com valores específicos na construção do discurso, tais quais como operador conclusivo, operador consecutivo e operador explicativo (MELO, BARBOSA, 2018).

Assim, como consequência da diversificação das categorias semântico-cognitivas que se associam ao uso do pronome, é de nosso interesse ampliar a investigação, enfatizando, neste momento, uma perspectiva onomasiológica. Após mapear os usos de **onde** na fala paulista (ALMEIDA, BERLINCK, 2019; ALMEIDA, 2019; ALMEIDA, 2020) e identificar um quadro de multifuncionalidade do pronome, o que aponta para uma possível mudança linguística associada ao item, devido ao processo de gramaticalização averiguado, temos o intento de observar a variação linguística no domínio de origem, isto é, no valor locativo. Desse modo, nosso estudo tem como objetivo investigar o processo de variação linguística das estratégias de relativização preposicionadas que expressam valor locativo na fala paulista.

Tal abordagem do fenômeno se justifica por meio do problema do encaixamento proposto por Weinreich, Labov e Herzog (2006 [1968], p. 132), haja vista que

A mudança linguística, ela mesma, raramente é um movimento de um sistema inteiro para outro. Em vez disso, descobrimos que um conjunto limitado de variáveis num sistema altera seus valores modais gradualmente de um polo para outro. As variantes das variáveis podem ser *contínuas* ou *discretas*; em qualquer dos casos, a variável mesma tem um espectro contínuo de valores, já que ele inclui a frequência de ocorrência de variantes individuais na fala estendida. O conceito da variável como um elemento estrutural torna desnecessário ver flutuações no uso como externas ao sistema, pois o controle de tal variação faz parte da competência linguística dos membros da comunidade de fala.

Dessa forma, argumentamos que a descrição da variação no domínio de origem de **onde** é um importante passo dentro da investigação maior dos usos do pronome para melhor compreensão da alteração que está em desenvolvimento dentro do sistema. Isso, pois nos permite identificar se há uma influência da multifuncionalidade, no sentido de abrir espaço para outras estratégias assumirem maior frequência no valor locativo, pensando nos valores modais que gradualmente se deslocam de um polo a outro. Nossa hipótese é de que ainda não nos encontramos nesse estágio da variação, dado que o valor locativo ainda é prototípico de **onde**. Contudo, temos duas razões para o presente estudo: (i) a próxima fase do trabalho tem por base

a diacronia, o que torna relevante analisar um corpus sincrônico para articular uma argumentação bem fundamentada a respeito de um possível processo de mudança; (ii) não há estudos que foquem na descrição do contexto sintático-semântico aqui especificado, o que nos impede de fazer uma afirmação categórica a respeito dos processos de variação e mudança e torna necessária a descrição.

Por isso, com base na bibliografia recente sobre as estratégias de relativização no Português Brasileiro (AMARAL, 1920; MOLLICA, 1977; LEMLE, 1978; TARALLO, 1983; CORRÊA, 1998; BARROS, 2000; BURGOS, 2003; BISPO, 2003; 2009), intencionamos averiguar se, há um espaço propício para a variação de formas de expressão do valor prototípico de lugar. Ademais, temos por objetivos:

- i. Identificar as estratégias de relativização que funcionam como variantes na expressão do valor locativo;
- ii. Determinar quais fatores linguísticos e extralinguísticos incidem sobre os usos dessas estratégias de relativização variantes;
- iii. Analisar em que medida a situação de variação no contexto locativo se articula com a multifuncionalidade de **onde** na expressão de outras categorias semântico-cognitivas;

Para que isso seja possível, nosso envelope de variação se constrói a partir das estratégias de relativização locativas do Português Brasileiro em contextos preposicionados, sendo elas:

Tabela 1 – Descrição do envelope de variação³

Estratégia	Definição	Exemplo
Padrão <i> pied- piping</i>	Em contextos preposicionados, a tradição gramatical prescreve o uso de preposição antes do pronome relativo.	Eu vou descrever a casa em que eu estou morando
Padrão com onde	A tradição gramatical ressalva que pronomes como onde não precisam da preposição, mesmo em contextos preposicionados.	Eu vou descrever a casa Ø onde eu estou morando
Copiadora	Caracterizada pelo pronome cópia dentro da oração relativa que retoma o antecedente.	Eu vou descrever a casa que eu estou morando lá/nela
Cortadora	Caracterizada pelo corte e/ou apagamento da preposição antes do pronome relativo.	Eu vou descrever a casa Ø que eu estou morando

Fonte: própria

O presente estudo foi detalhado a partir de 5 seções, sendo a segunda, *Fundamentação teórica*, dividida em três grandes frentes: *Por onde anda você? Uma breve contextualização*, que trata dos principais resultados encontrados na perspectiva semasiológica do mapeamento de **onde**, assim como outros estudos; *O objeto de estudo*, responsável por introduzir as estratégias de relativização com que estamos trabalhando, a partir de duas perspectivas diferentes: (i) uma descritiva, tendo como objetivo destrinchar as considerações da tradição normativa e descritiva a respeito das orações relativas, e (ii) um *overview* linguístico, explorando as análises de trabalhos sobre esse fenômeno no Português Brasileiro; por fim, *Teoria de Variação e Mudança Linguística*, com o intento de tecer considerações a respeito do nosso suporte teórico-metodológico, ressaltando as questões que concernem as variáveis sintáticas. A terceira seção, *Decisões teórico-metodológicas*, apresenta como foi construído o *corpus* utilizado no estudo, assim como considerações relevantes sobre a análise dos dados, desde sua “filtragem” até as variáveis independentes selecionadas. A seção subsequente, *A sintaxe dos paulistas: análise de produção*, se destina à descrição dos resultados de produção; enquanto a sétima seção aponta as considerações finais da investigação.

³ Para o contexto preposicionado, com valor locativo, não encontramos usos de *preposition stranding* ou *pied-piping* (padrão com preposição) encabeçadas por **a qual, na qual**. Para mais detalhes do envelope de variação, ver seção 3: *Decisões teórico-metodológicas*.

5 CONSIDERAÇÕES FINAIS

O presente estudo tinha como objetivo principal averiguar a variação de formas na expressão do valor prototípico de lugar. Para que isso fosse possível, mobilizamos um envelope de variação com as estratégias de relativização locativas em contextos preposicionados na fala paulista, obtendo, assim, as variantes: **padrão com preposição** (*pedpiping*), **padrão com onde**, **copiadora** e **cortadora**.

Tendo em vista que esse momento da pesquisa se insere em uma investigação maior com o objetivo de compreender os usos de **onde**, foi necessário, na seção 2, apresentar uma breve contextualização a respeito de passos anteriores da pesquisa. Além disso, consideramos adequado trazer algumas provocações provenientes de uma perspectiva interlinguística, articulando como o processo que está ocorrendo com nosso pronome relativo locativo foi observado em outros idiomas.

Ainda nessa seção, apresentamos uma breve fundamentação teórica a respeito das estratégias de relativização, ressaltando como nosso fenômeno se justifica pela falta de referências a respeito do contexto preposicionado locativo no Português Brasileiro. Nosso enfoque perpassou os trabalhos que discutiam o impacto das variáveis extralinguísticas no *corpus* tendo em vista a natureza de nossa análise. Por último, apresentamos nosso apoio teórico-metodológico, a Teoria de Variação e Mudança Linguística, assim como alguns problemas condizentes ao trabalho com variáveis sintáticas.

Sobre esse último, elencamos três principais problemáticas: (i) a baixa frequência de dados; (ii) a falta de correlação com fatores sociais; e (iii) a necessidade de suporte teórico para a análise. Nesse sentido, nosso trabalho apresenta uma quebra, visto que temos um número expressivo de dados (1007 no total) e as variáveis extralinguísticas se mostraram mais fortes na explicação do fenômeno em variação.

Dando continuidade ao trabalho desenvolvido, na seção 3 apresentamos nossas decisões teórico-metodológicas. Assim, descrevemos os bancos de dados com os quais trabalhamos, sendo eles, o banco de dados IBORUNA do projeto ALIP (Gonçalves, 2007), representando a fala do interior paulista, e o banco de dados Projeto SP2010 (MENDES, 2013), representando a fala da capital paulistana. Evidenciamos as diferenças presentes na constituição e objetivo dos bancos de dados que foram extremamente importantes para nossa análise.

Ademais, descrevemos as variáveis linguísticas que controlamos em nosso *corpus*, sendo elas: (i) função sintática do pronome relativo; (ii) adjacência entre o sintagma nominal

antecedente e a oração relativa; (iii) classe morfológica do locativo; (iv) identidade semântica (ou tipo) do locativo; (v) classificação semântica da relativa (em explicativa ou restritiva); (vi) definitude do antecedente; (vii) especificidade do antecedente; e (viii) status informacional do antecedente. Reforçamos, também, as variáveis extralinguísticas provenientes dos bancos de dados: (i) sexo/gênero; (ii) escolaridade; (iii) faixa etária (presentes nos dois bancos); (iv) tipologia textual (presente no IBORUNA); (v) região de residência de São Paulo; (vi) zona de residência de São Paulo; (vii) roteiro da entrevista (presentes no Projeto SP2010).

Na seção 4, subdividimos a nossa análise em três momentos: (i) a discussão da análise de regressão logística de efeitos mistos; (ii) a descrição da análise univariada; e, por fim, (iii) a exposição de algumas considerações a respeito das estratégias copiadora e padrão com preposição.

Isso se deu pela amostra com que estávamos trabalhando, já que, para o banco de dados IBORUNA, obtivemos 3 dados **de padrão com preposição**, 35 dados de **copiadora**, 249 dados de **cortadora** e 267 dados de **padrão com onde**; para o banco de dados Projeto SP2010, obtivemos 10 dados de **padrão com preposição**, 15 dados de **copiadora**, 128 dados de **padrão com onde** e 300 dados de **cortadora**. Assim, produzimos uma análise estatística consistente das estratégias **padrão com onde** e **cortadora**, para os dois bancos de dados, que privilegiou a análise multivariada. Nossa intenção era compreender como cada uma das variáveis atuavam na explicação do fenômeno em variação. Contudo, achamos coerente também apresentar os resultados que não haviam obtido significância estatística, projetando, principalmente, a replicabilidade da pesquisa.

Desse modo, averiguamos, logo de início, que as diferenças na constituição dos bancos de dados mobilizaram contextos sintáticos, semânticos e pragmáticos diferentes. Ao explorar a identidade lexical do verbo da oração relativa, observamos frequências e itens distintos, o que impactou nos demais fatores linguísticos controlados. Assim, para o modelo de regressão, nos dois *corpora*, tínhamos como variável aleatória o falante e a identidade lexical e como valor de aplicação da variável resposta as estratégias de relativização **padrão com onde**. Como variáveis previsoras, tínhamos, para o banco de dados IBORUNA: faixa etária, escolaridade e identidade semântica do locativo; para o banco de dados Projeto SP2010: faixa etária, região do falante, escolaridade e definitude do antecedente.

Resumidamente, concluímos que, para o banco de dados IBORUNA, há um desfavorecimento do uso da **padrão com onde** para falantes da primeira faixa etária em comparação com o *intercept*. Tal diferença entre as estimativas não é significativa, mas essa faixa é responsável pelo maior uso da **cortadora**, enquanto a segunda é responsável pelo maior

uso da **padrão com onde**, o que confirma os dois movimentos de “picos” relacionando as faixas etárias com usos de estratégias diferentes. Outra diferença interessante, e que reforçou nossa análise, foi o desfavorecimento do uso da variante padrão pelos níveis de escolaridade mais baixos, em especial o ensino fundamental. Isso aponta para uma relação do uso da **padrão com onde** com o contato com a norma culta durante os anos escolares. Por fim, observamos o desfavorecimento da estratégia quando o antecedente se refere a um local.

Para o banco de dados do projeto SP2010 há um desfavorecimento do uso da **padrão com onde** para falantes da segunda e terceira faixa etária em comparação com o *intercept*. Tal diferença entre as estimativas não é significativa, estatisticamente ($p > 0.05$), porém essas faixas são responsáveis pelo maior uso da **cortadora**. A primeira faixa etária é responsável pela frequência mais alta da estratégia **padrão com onde** no *corpus*, o que confirma os dois movimentos de “picos” relacionando às faixas etárias com usos de estratégias diferentes. Também observamos um desfavorecimento do uso da variante padrão pelo nível de escolaridade mais baixo, ou seja, o ensino médio. Isso aponta, tal qual para o banco de dados IBORUNA, para uma relação do uso da **padrão com onde** com a norma. Por fim, observamos o desfavorecimento da estratégia para moradores da região centro velho e em contextos indefinidos.

Dessa forma, discutimos que os fatores extralinguísticos, por serem semelhantes nos dois bancos de dados, apresentam um poder explanatório maior no processo de variação das estratégias de relativização locativas em contextos preposicionados na fala paulista. Assim, a idade e escolaridade do falante são perspectivas interessantes para estudos futuros compreenderem o significado social que as variantes assumem na fala dos paulistas.

A análise univariada nos revelou que a adjacência entre o sintagma nominal antecedente e a oração relativa, a classificação semântica da relativa, a classe morfológica do locativo, a função sintática do pronome relativo, o sexo/gênero do falante, a tipologia textual e o roteiro de entrevista não apresentam diferenças estatisticamente significantes entre os níveis de análise. Além disso, em alguns casos, a não significância decorre de um problema de ortogonalidade, de um desbalanceamento dos dados. No caso da adjacência, por exemplo, vemos uma tendência comum aos dois bancos que poderia se confirmar se houvesse um conjunto mais robusto de dados de tipo ‘distante’. Assim, esses fatores não contribuíram, no contexto de nosso estudo, para a compreensão do fenômeno.

Por fim, com relação à **copiadora** e à **padrão com preposição**, consideramos adequado discutir os fatores envolvidos em sua baixa incidência no *corpus*. A partir da literatura, concluímos que o valor negativo associado a essa estratégia não padrão e a falta de naturalidade

da estratégia padrão são os responsáveis pela pouca frequência. Outrossim, identificamos adjacência entre o sintagma nominal antecedente como um contribuinte para o uso dessas variantes, já que contextos antecedentes exigem menos esforço cognitivo do falante. A escolaridade também influencia na escolha de uma ou outra, tendo em vista que falantes mais escolarizados são os responsáveis por produzir a **padrão com preposição**, enquanto os menos escolarizados utilizam a **copiadora**.

Após as explicações teórico-metodológicas e a investigação estatística, concluímos que existe um processo de variação linguística na expressão do valor locativo em contextos preposicionados, por meio das variantes: **padrão com preposição**, **padrão com onde**, **cortadora** e **copiadora**. A partir dos fatores linguísticos e extralinguísticos que condicionam o uso das estratégias, constatamos que afala do interior paulista se utiliza, em maior frequência, do uso prototípico do pronome **onde**; enquanto a fala da capital paulistana apresenta um uso maior da cortadora, que já havia sido identificado em momentos anteriores (TARALLO, 1983).

De todo modo, essa etapa da investigação era necessária considerando o problema do encaixamento da Teoria de Variação e Mudança Linguística. Assim, tendo em vista que o processo de mudança linguística não é um movimento de deslocamento, mas sim um processo em que “um conjunto limitado de variáveis num sistema altera seus valores modais gradualmente de um polo para outro” (WEINREICH, LABOV, HERZOG, 2006 [1968], p. 132), identificamos como o domínio de origem do pronome relativo **onde** se encontra apesar do processo de gramaticalização. Dessa forma, verificamos um processo de variação entre as estratégias relativas.

Como encaminhamentos futuros, nos dedicaremos a uma análise diacrônica para avaliarmos, com base nos estágios da língua, o desenvolvimento da multifuncionalidade de **onde** e o impacto desse desenvolvimento nas estratégias de relativização locativas em contexto preposicionado no Português Brasileiro. Isso nos dará os recursos necessários para afirmações mais categóricas a respeito dos processos que estamos investigando.

REFERÊNCIAS

- ALMEIDA, M. A. de; BERLINCK, R. de A. 2019. “Por onde anda você?” – Sobre a norma e o uso de onde na fala paulista. **Revista Diálogos**, v. 7, n. 1, p. 70-88, 2019.
- ALMEIDA, M. A. de. **Onde anda você?** um estudo sobre os usos de “onde” na fala paulista. Araraquara: Trabalho de conclusão de curso (bacharelado - Letras) - Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências e Letras (Campus de Araraquara), 2019.
- ALMEIDA, M. A. de. “Onde está você agora?” – Uma investigação sobre os contextos não locativos de onde na fala paulista. **Revista Falange Miúda**, v. 5, n. 2, p. 157-175, 2020.
- AMARAL, A. **O dialeto caipira**. São Paulo: Hucitec-SCET-CEC, 1976 [1920].
- ARAÚJO, L. S. de. **A expressão dos valores 'antepresente' e 'passado absoluto' no espanhol**: um olhar atento a variedades diatópicas da Argentina e da Espanha. Orientador: Rosane de Andrade Berlinck. 2017. F. Tese (Doutorado em Linguística e Língua Portuguesa) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, São Paulo, 2017. Disponível em: < <https://repositorio.unesp.br/handle/11449/150649?show=full> > Acesso em: 01 março 2020.
- ARIEL, M. Referring and accessibility. **J. Linguistics**, 24, p. 65-87, 1988.
- ARIEL, M. The function of accessibility in a theory of grammar. **Journal of Pragmatics**, 16, p. 443-463, 1991.
- BARROS, A. L. de. **O uso da relativa cortadora na fala pessoense**. Dissertação de Mestrado orientada pelo professor Dermeval da Hora, UFPB, João Pessoa, 2000.
- BEAMAN, Karen V. Swabian relatives: variation in the use of the wo-relativiser. In: **Advancing Socio-grammatical Variation and Change**. Routledge, 2021. p. 134-164.
- BECHARA, E. **Moderna gramática portuguesa**. 37 ed. Rio de Janeiro: Nova Fronteira, 2009.
- BERLINCK, R.A. **La position du sujet en portugais**. Étude diachronique des variétés brésilienne et européenne. Thèse de Doctorat. Katholieke Universiteit Leuven, 1995.
- BISPO, E. B. **Relativa copiadora**: uso, regularização e ensino. Dissertação de Mestrado orientada por Maria Angélica Furtado da Cunha. Programa de Pós-Graduação em Estudos da Linguagem, UFRN, Natal, 2003.
- BISPO, E. B. **Estratégias de relativização no português brasileiro e implicações para o ensino**: o caso das cortadoras. 164f. Tese (Doutorado em Estudos da Linguagem), UFRN, Natal-RN, 2009.
- BOURDIEU, P. L'économie des échanges linguistiques. **Langue Française**, Paris, n.34, 1977.
- BOURDIEU, P. Price Formation and the Anticipation of Profits. In: BOURDIEU, P. **Language and Symbolic Power**. Cambridge: Polity Press, 1991.

BRAGA, M.; MANFILI, K. Essa é a preocupação onde eu quero chegar: “onde” em referências anafóricas no português do Brasil. **Veredas** 8 (1-2), 2004. p.233-243.

BRAGA, M. L.; ILARI, R.; OLIVEIRA, R. P. de; BASSO, R. M. Os especificadores – parte 1: artigo definido. In: ILARI, R. (org.). **Palavras de classe fechada** Gramática do português culto falado no Brasil, v. 4). São Paulo: Contexto, 2015, p. 107-128.

BRANDÃO, S.M. Mercado linguístico: uma interpretação da imbricada relação estrutura linguística e estrutura social. **Linguagem: Estudos e Pesquisas**. Catalão-GO. V. 21, n.1, 2017. p.225-255.

BRANDÃO, S. M. **Alternância verbal em construções condicionais - um fenômeno variável?**. Araraquara: Dissertação (mestrado) - Universidade Estadual Paulista “Júlio de Mesquita Filho”, Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Letras, Araraquara, 2018.

BRANDÃO, S. M. **Template - análises linguísticas multivariadas** (variáveis binomiais) no R. [S.l.]: Zenodo. , 30 Sep. 2018. DOI: <https://doi.org/10.5281/zenodo.1439457>

BROOK, Marisa. One of those situations where a relative pronoun becomes a complementizer: A case of grammaticalization in progress... again. In: **Proceedings of the 2011 Canadian Linguistics Association annual conference**. 2011.

BURGOS, L.E.S.de. **Estratégias de uso das relativas em uma comunidade de fala afro-brasileira**. Dissertação de Mestrado, Universidade Federal da Bahia, 2003.

CALDEIRA, T. P. do R. **Cidade de muros: crime, segregação e cidadania em São Paulo**. 3 ed. São Paulo: Editora 34, 2011.

CAMBRAIA, C.N; ARAÚJO, L.E.S. Variação em locativos no português de Belo Horizonte: estudo sociolinguístico. **Paralelo** 20, Belo Horizonte, n. 2, dez. 2004. p.123-132.

CASTILHO, A.T. de. **Nova gramática do português brasileiro**. São Paulo: Contexto, 2014.

CORRÊA, W.R. **Oração relativa: o que se fala e o que se aprende no português do Brasil**. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas, 1998.

CUNHA, C.; CINTRA, L. **Nova gramática do português contemporâneo**. 5. ed. Rio de Janeiro: Lexikon, 2008.

FREITAG, Raquel Meister Ko. Problemas teórico-metodológicos para o estudo da variação linguística nos níveis gramaticais mais altos. **Matraga - Revista do Programa de Pós-Graduação em Letras da UERJ**, [S.l.], v. 16, n. 24, dez. 2009. ISSN 2446-6905. Disponível em: <<https://www.e-publicacoes.uerj.br/index.php/matraca/article/view/27799/19920>>. Acesso em: 29 jul. 2021.

FREITAG, R. M. K. (Re)discutindo sexo/gênero na sociolinguística. In: FREITAG, R. M. K.; SEVERO, C. G. (org). **Mulheres, linguagem e poder: estudos de gênero na sociolinguística brasileira**. São Paulo: Blucher, 2015, p. 17-73.

FREITAG, R. M. K. **Teste de hipóteses e valores de significância**. Disponível em: <https://rkofreitag.github.io/Pvalor.html/>. Acesso em: 2020-10-07.

FREITAG, R. M. K. **Variáveis categóricas**. Disponível em: <https://rkofreitag.github.io/Categorica.html/>. Atualizado em: 2021-04-11.

FREITAG, R. M. K. **Redução de variáveis e de dados**. Disponível em: <https://rkofreitag.github.io/reducao.html/>. Atualizado em: 2021-04-11.

HEINE, B.; CLAUDI, U.; HÜNNEMEYER, F. **Grammaticalization: a Conceptual Framework**. Chicago: The University of Chicago Press, 1991.

HOPPER, P. J.; TRAUGOTT, E. C. **Grammaticalization**. New York, Cambridge University Press, 1993.

HYNDMAN et al (2021). **forecast**: Forecasting functions for time series and linear models. Disponível em: < <https://pkg.robjhyndman.com/forecast/> >

GILI GAYA, Samuel. **Curso superior de sintaxis española**. 9 ed. Barcelona: Biblograf, 1970.

GONÇALVES, Sebastião Carlos Leite. **Banco de dados Iboruna**: amostras eletrônicas do português falado no interior paulista. Disponível em: <http://www.alip.ibilce.unesp.br>.

KATO, M. A. Orações relativas: variação universal e variação individual no português. **Estudos lingüísticos**, V, 1981, p. 1-16.

KATO, M. Recontando a história das relativas em uma perspectiva paramétrica. In: ROBERTS, I.; KATO, M. **Português brasileiro**: uma viagem diacrônica. Campinas: Editora da Unicamp, 1996.

LABOV, W. **Padrões Sociolinguísticos**. São Paulo: Parábola Editorial, 2008 [1972].

LABOV, W. Where does the linguistic variable stop? A response to Beatriz Lavandera. **Working Papers in Sociolinguistics**, 44. Austin: Southwest Educational Development Laboratory, 1978.

LABOV, W. **Principles of Linguistic Change**. Vol. 1: Internal Factors. Cambridge, MA/Oxford: Blackwell Publishers. 1994

LABOV, W. **Principles of Linguistic Change**. Vol. 2: Social Factors. Cambridge, MA/Oxford: Blackwell Publishers. 2001.

LABOV, W. Some Sociolinguistic Principles. In: PAULSTON, C.B.; TUCKER, G.R. (eds) **Sociolinguistics: The Essential Readings**. Oxford: Blackwell Publishing, 2003. p.234-250.

LABOV, W. **Principles of Linguistic Change**. Vol. 3: Cognitive Factors. Cambridge, MA/Oxford: Blackwell Publishers. 2010.

LABOV, W. **The Social Stratification of English in New York City**. Washington: Center for Applied Linguistics, 1966.

LAKOFF, G.; JOHNSON, M. *Metaphor we live by*. Chicago: The University of Chicago Press, 1980.

LAVANDERA, B. Where does the sociolinguistic variable stop?. **Language in Society**, v.7, p. 171-182, 1978.

LEHMANN, C. **Thoughts on grammaticalization**. 2 ed. Erfurt: Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfut, 2002.

LEMLE, M. “Heterogeneidade dialectal: uma apelo à pesquisa” In *Linguística e Ensino do Vernáculo*, organizado por Lobato, **Tempo Brasileiro**, 53/53: 60-94, 1978.

LEVSHINA, N. **How to do Linguistics with R**. Data exploration and statistical analysis. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2015.

LIMA, S.C. **O impacto do vernáculo sobre o uso do onde na escrita monitorada**. 2007. 142 f. Dissertação (Mestrado em Linguística) – Instituto de Letras, Universidade de Brasília, Distrito Federal. 2007.

LIMA, S.; COROA, M.L.M.S. Recategorização, semantização e discursivização na trajetória de gramaticalização do onde. **Via Litterae**, Anápolis, v. 5, n. 2, p. 307-333, jul./dez, 2013. Disponível em: <http://www2.unucseh.ueg.br/vialitterae>

MALLMANN, A. C. L. G.; GARCIA, J. C. V.; ESCOBAR, R. de C. P. S. Variação em estratégias de relativização no Português Brasileiro. VIEIRA, S. R.; LIMA, M. D. A. de O. (orgs.). **Variação, gêneros textuais e ensino de Português: da norma culta à norma-padrão**. Rio de Janeiro: Letras UFRJ, 2019, p. 63-88.

MARINHO, J. H. C. **O uso do onde no texto acadêmico**. Revista Estudos Linguísticos, Belo Horizonte, v.8, n.1, p.159-170, jan./jun. 1999.

MARTELOTTA, M. E. **Mudança linguística: uma abordagem baseada no uso**. São Paulo: Cortez, 2011.

MATEUS, M.H.M et al. **Gramática da Língua Portuguesa**. 6 ed. Lisboa: Editorial Caminho, 2003.

MELO, R. M.; FRANCA, J. C. O item onde e suas rotas de mudança: uma abordagem à luz da teoria da gramaticalização. In: II Congresso Brasileiro sobre Letramentos e Dificuldades de Aprendizagem (CONBRALE), 2018, CAMPINA GRANDE. **Anais II CONBRALE. CAMPINA GRANDE**: Realize, 2018. v. 1. p. 1-10.

MENDES, R.B. **Projeto SP2010: Amostra da fala paulistana**. 2003. Disponível em <<http://projetosp2010.fflch.usp.br>>. Acesso em 13/04/2019.

MILROY, L.; GORDON, M. Beyond Phonology: Analyzing and Interpreting Higher Lever Variation. In: MILROY, L.; GORDON, M. **Sociolinguistics: Method and Interpretation**. Oxford: Blackwell Publishing, 2003.

MOLLICA, M. C. **Estudo da cópia nas construções relativas em português**. Rio de Janeiro, 1977. Dissertação (Mestrado) Pontifícia Universidade Católica.

MOLLICA, M. C. **Influência de fatores de processamento na variação em português**. In: TARALLO, F. (org). **Fotografias sociolinguísticas**. Campinas: Pontes, 1989.

MOLLICA, M. Relativas em tempo real no português brasileiro contemporâneo. In: PAIVA, M. C.; DUARTE, M. E. **Mudanças linguísticas em tempo real**. Rio de Janeiro: Contra Capa, 2003, p.129-138.

NEVES, M.H.M.; BRAGA, M.L. As construções hipotáticas/adverbiais. In: NEVES, M.H.M. (org.) **A construção das orações complexas** (Gramática do português culto falado no Brasil, v. 5). São Paulo: Contexto, 2016. p.124-166.

NEVES, M. H. de M. **A gramática do português revelada em textos**. São Paulo: Editora Unesp, 2018.

NICHOLAS, N. **The story of pu**. Melbourne: The University of Melbourne dissertation. 1998

OLIVEIRA e SILVA, G.M.; PAIVA, M.C. Visão de conjunto das variáveis sociais. In: OLIVEIRA e SILVA, G.M.; SCHEREE, M.M.P. **Padrões Sociolinguísticos: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro**. Rio de Janeiro: Tempo Brasileiro, 1996. p.335-378.

OLIVEIRA, M. R. de. Categorias cognitivas em debate: a trajetória dos pronomes locativos em destaque. In: LIMA-HERNANDEZ, M.C. **Gramaticalização em perspectiva: cognição, textualidade e ensino**. São Paulo: Paulistana, 2010.

OUSHIRO, L. **Identidade na pluralidade: avaliação, produção e percepção linguística na cidade de São Paulo**. 2015. Tese (Doutorado em Semiótica e Lingüística Geral) - Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2015.

OUSHIRO, L. (2018) **dmsocio**. v0.2.0. Available at <oushiro.shinyapps.io/dmsocio>

OUSHIRO, L. (2021). Introdução à Estatística para Linguistas (1.0.9). Zenodo. <https://doi.org/10.5281/zenodo.4584104>

PATIL, I. (2018). **ggstatsplot: "ggplot2" Based Plots with Statistical Details**. CRAN. Retrieved from <<https://cran.r-project.org/web/packages/ggstatsplot/index.html>>

PINHEIRO, C. L. **A relativização no português oral culto de Fortaleza**. Dissertação de Mestrado – Universidade Federal do Ceará, 1998.

PRINCE. E. F. Toward a Taxonomy of Given-New Information. In: COLE, P. (ed.). **Radical pragmatics**. New York: Academic Press, 1981, p. 223-233.

RAE. **Esbozo de una nueva gramática de la lengua española**. Madrid: Espasa, 1986.

R Core Team (2021). **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2018. Disponível em: <https://www.R-project.org/>.

ROCHA LIMA, C. H. **Gramática normativa da língua portuguesa**. Rio de Janeiro: José Olympio, 2020 [1972].

ROUSSOU, A. Complements, relatives, and nominal properties. **Handout from talk presented at the University of Ghent**, 2012.

SAID ALI, M. **Gramática histórica da língua portuguesa**. 8. ed. Brasília: Companhia Melhoramentos, 2001.

SANKOFF, G.; KEMP, W., CEDERGREN, H. The syntax of *ce que qu'est-ce que* variation and its social correlates. In: SHUY, R. E; FIRSCHMIG, R.W. (eds) **Dimensions of variability and competence**. Washington: Georgetown University Press, 1980.

SANTOS, Aymmé Silveira; MELO, Raniere Marques de. O ensino da variação linguística na Base Nacional Comum Curricular. **Entrepalavras**, Fortaleza, v. 9, n. 3, p. 115-132, setdez/2019.

SEVERO, C. G. O papel do gênero/sexo nos estudos sociolinguísticos de variação/mudança. **Revista de Letras**, v. 8, p. 1-08, 2006.

SOUZA, E. H. P. M. de. **A multifuncionalidade do onde na fala de Salvador**. 2003. Tese (Doutorado em Letras e Linguística) - Instituto de Letras, Universidade Federal da Bahia, Bahia.

SOUZA, I. G. de; CARVALHO, C. dos S. Usos de Onde na Escrita: Multifuncionalidade, Abstratização Semântica e Gramaticalização. In: LOPES, N. da S.; SANTOS, E. S. dos S.; CARVALHO, C. S. de. **Língua e Sociedade: Diferentes Perspectivas**, Fim Comum. São Paulo: Blucher, 2019, p. 197-214.

TANG, W. **A Simple Guide to Using Antconc**. 2011. Disponível em: <http://www.laurenceanthony.net/software/antconc/resources/help_AntConc321_english.pdf>

TARALLO, F.L. **Relativization strategies in Brazilian Portuguese**. Philadelphia: University of Pennsylvania, 1983. PhD Dissertation.

TARALLO, F.L. The filling of the gap PRO-DROP rules in Brazilian Portuguese. In: KING, L.D.; MALEY, C.A. (eds) **Selected papers from the XIIIth Linguistic Symposium on Romance Languages**. (Current Issues in Linguistic Theory, 36), Amsterdam/Philadelphia: John Benjamins, 1985.

TARALLO, F. L. **A pesquisa sociolinguística**. São Paulo: Editora Ática, 1986.

TORREGO, L. G. **Gramática didáctica del español**. 8 ed. Madrid: SM, 2002.

TRUDGILL, Peter. Sex, covert prestige and linguistic change in the urban British English of Norwich. **Language in society**, v. 1, n. 02, p. 179-195, 1972.

WEINREICH, U.; LABOV, W.; HERZOG, M. **Fundamentos empíricos para uma teoria da mudança linguística**. Trad. Marcos Bagno. São Paulo: Parábola, 2006 [1968]

WILMET, M. **La détermination nominale**. Quantification et caractérisation. Paris: Presses Universitaires de France, 1986.

ZILLES, A. M. S.; KERSCH, D. F. Onde: prescrição, proscricção, descrição e ensino. In: ZILLES, A. M. S.; FARACO, C. A. (Org.). **Pedagogia da Variação Linguística: língua, diversidade e ensino**. São Paulo: Parábola Editorial, 2015, p. 145-187.