

REVISÃO DE LITERATURA

META-ANÁLISE DOS ESTUDOS DE NEGAÇÃO VERBAL NAS REGIÕES NORDESTE E SUDESTE: INVESTIGAÇÃO DA RELEVÂNCIA DOS CONDICIONAMENTOS SOCIAIS



OPEN ACCESS

Todo conteúdo de *Cadernos de Linguística* está sob Licença Creative Commons CC - BY 4.0.

EDITORES

- Raquel Freitag (UFS)
- Juliana Bertucci (UFTM)
- Márcia Vieira (UFRJ)

AVALIADORES

- Paloma Cardoso (UFS)
- Manoel Siqueira (SEDUC/AL)

SOBRE OS AUTORES

- Pedro Henrique Sousa dos Santos
Conceitualização; Investigação;
Metodologia; Software; Análise Formal;
Visualização; Escrita - Rascunho Original.
- Elyne Giselle de Santana Lima Aguiar VITÓRIO
Conceitualização; Metodologia; Supervisão;
Validação; Escrita - Análise e Edição.

Recebido: 02/04/2025

Aceito: 22/07/2025

Publicado: 22/12/2025

COMO CITAR

SANTOS, P.H.S.; VITÓRIO, E.G.S.L.A. (2025). Meta-análise dos estudos de negação verbal nas Regiões Nordeste e Sudeste: investigação da relevância dos condicionamentos sociais. *Cadernos de Linguística*, v. 6, n. 4, e858.



VERIFICAR
ATUALIZAÇÕES

Pedro Henrique Sousa dos SANTOS

Programa de Pós-Graduação em Linguística - Universidade Federal do Rio de Janeiro (UFRJ)

Rio de Janeiro, Rio de Janeiro, Brasil

Elyne Giselle de Santana Lima Aguiar VITÓRIO

Faculdade de Letras - Universidade Federal de Alagoas (UFAL)

Maceió, Alagoas, Brasil

RESUMO

O estudo de fenômenos linguísticos variáveis é uma área de pesquisa em constante expansão no Brasil, principalmente dentro do campo da Sociolinguística Variacionista (Labov, 2008). Entretanto, com a profusão de trabalhos nos mais diversos níveis gramaticais, a Sociolinguística passa a necessitar de métodos de sistematização das pesquisas sobre cada tema a fim de se encontrar possíveis lacunas e tecer generalizações. Entre eles, a revisão sistemática com meta-análise é considerada a com maior poder explicativo (Freitag, 2021). Nesse contexto, tomando como base a negação verbal, este trabalho objetivou investigar a relevância de três condicionamentos sociais em uma meta-análise, a saber: sexo/gênero, escolaridade e faixa etária. Para isso, foi preciso delimitar a análise a duas regiões, Nordeste e Sudeste, por conta da inter-relação histórica entre elas (Villa, 2017). Como resultados desta meta-análise, foram percebidas baixa ou nula influência de "sexo/gênero", significante influência de "escolaridade" e baixa ou nula influência de "faixa etária". Nesse

sentido, o trabalho contribuiu para a ampliação de pesquisas que versem sobre meta-análise na Sociolinguística Variacionista, apontando para meios que precisam ser padronizados a fim de que a comparação entre os estudos de regiões urbanas, principalmente, seja melhor executada e os resultados sejam melhor sistematizados.

PALAVRAS-CHAVE

Sociolinguística; Variação; Negação Verbal; Meta-Análise;
Condicionamento Social.

TITLE

META-ANALYSIS OF VERBAL NEGATION STUDIES IN THE NORTHEAST
AND SOUTHEAST REGIONS: INVESTIGATION OF THE RELEVANCE OF
SOCIAL CONDITIONINGS

ABSTRACT

The study of variable linguistic phenomena is a constantly expanding area of research in Brazil, mainly within the field of Variationist Sociolinguistics (Labov, 2008). However, with the profusion of studies at the most diverse grammatical levels, Sociolinguistics needs methods for systematizing research on each topic in order to find possible gaps and make generalizations. Among them, the systematic review with meta-analysis is considered the one with the greatest explanatory power (Freitag, 2021). In this context, based on verbal negation, this work aimed to investigate the relevance of three social conditionings in a meta-analysis, namely: sex/gender, education and age group. To do this, it was necessary to limit the analysis to two regions, Northeast and Southeast, due to the historical interrelationship between them (Villa, 2017). As results of this meta-analysis, we found low or no influence of "sex/gender", a significant influence of "education" and low or no influence of "age group". Therefore, the work contributed to the expansion of works that deal with meta-analysis in Variationist Sociolinguistics, pointing to elements that need to be standardized so that the comparison between studies of urban regions, mainly, is better executed and the results are better systematized.

KEYWORDS

Sociolinguistics; Variation; Verbal Negation; Meta-Analysis;
Social Restriction.

INTRODUÇÃO

A Sociolinguística Variacionista (Labov, 2008) trata da variação como um fenômeno condicionado por restrições linguísticas e sociais. Contudo, diferentes tipos de fenômenos são influenciados distintamente pelas variáveis sociais. Em geral, as variáveis do nível fonético-fonológico tendem a ser mais afetadas, dado que não possuem significado referencial (Eckert; Labov, 2017). No caso das variáveis morfossintáticas, a situação é mais problemática, dado que elas carregam significado referencial, o que gerou uma relevante discussão na história da Sociolinguística entre Beatriz Lavandera (Lavandera, 1978) e William Labov (Labov, 1978), do qual se concluiu que, mesmo em casos nos quais não há condicionamento social, o estudo da variação morfossintática ainda auxilia no entendimento da gramática da língua.

Os estudos que enfocam a produção na Sociolinguística – isto é, que usam entrevistas para elicitare narrativas pessoais dos informantes – utilizam variáveis sociais como escolaridade, sexo, gênero, idade ou faixa etária independentemente do nível da variação, se fonológica, morfológica, sintática etc. Em alguns casos, a variação nos níveis gramaticais mais altos – no sentido proposto por Eckert e Labov (2017) – ainda pode revelar influência de condicionantes sociais, como no caso do (ING) no inglês (Campbell-Kibler, 2009) e da concordância verbal de terceira pessoa do plural (Almeida, 2022). Paralelamente, há fenômenos que não parecem ser significativamente influenciados por variáveis sociais – especialmente os mais frequentemente utilizados na Sociolinguística brasileira, a saber: sexo/gênero, escolaridade e faixa etária (Mollica; Braga, 2020) –, como apontam Novais e Siqueira (2020). Nesse sentido, é importante haver, na Sociolinguística, espaço para pesquisas que avaliem o quanto informações sociais são importantes para a explicação de determinado fenômeno linguístico, especialmente no campo morfossintático.

O presente estudo, nessa linha, tem como objetivo identificar o efeito dos condicionamentos sociais no uso de um fenômeno morfossintático por meio de uma revisão sistemática com meta-análise, uma técnica de análise de revisões que agrupa os dados dos estudos e aplica a eles os mesmos testes estatísticos. Como esse procedimento é de recente implementação na área da Linguística, a pesquisa justifica-se por ampliá-lo, visto que a meta-análise auxilia na comparação dos resultados dos trabalhos por métodos estatísticos homogêneos de modo a eliminar possíveis vieses de revisões narrativas e ampliar o poder explanatório da revisão (Freitag, 2021).

Para isso, foi selecionado o fenômeno da negação verbal, que é constituído pelas variantes “Negação Pré-verbal”, “Negação Pós-verbal” e “Dupla Negação”, como (1a), (1b) e (1c)¹, respectivamente. Como recorte de pesquisa, são selecionadas as regiões Nordeste e Sudeste pelo

¹ Exemplos criados com base em introspecção.

fato de as duas regiões apresentarem inter-relação histórica devido às migrações humanas durante o século XX (Villa, 2017). Assim, pretende-se investigar se a negação é influenciada pelas variáveis sociais sexo/gênero, faixa etária e escolaridade, usualmente escolhidas para estratificação de amostras no Brasil.

- (1) a. A menina *não* gosta de mim.
b. A menina gosta de mim *não*.
c. A menina *não* gosta de mim *não*.

Nesse sentido, foram elencados os seguintes objetivos específicos: i) realizar uma revisão sistemática da literatura sobre o fenômeno abordado; ii) contrastar diferentes resultados para o comportamento dos fatores sociais sexo/gênero, escolaridade e faixa etária; iii) discutir o efeito social das variáveis sobre a distribuição das variantes da negação verbal; e v) socializar fatores sociais significativos para a pesquisa da negação.

Para esses objetivos, elencam-se as seguintes perguntas norteadoras: i) o uso da negação verbal difere-se nas regiões Nordeste e Sudeste?; e ii) as tradicionais restrições extralinguísticas exercem pressão sobre a variação na negação verbal? Como respostas provisórias a esses questionamentos, dispõem-se duas hipóteses principais: i) há um comportamento distinto entre as regiões Nordeste e Sudeste, principalmente no que se refere ao maior uso de Dupla Negação e Negação Pós-verbal pelos falantes nordestinos; e ii) acerca das variáveis analisadas na meta-análise, não haverá condicionamento das variáveis por ser a negação um fenômeno morfossintático, geralmente mais abaixo do nível de consciência social.

Assim, o trabalho está organizado da seguinte maneira: na seção 1, sintetiza-se o fenômeno da negação no português brasileiro. Na seção seguinte, destaca-se a fundamentação teórica para o trabalho, enquanto a quarta seção descreve as ferramentas estatísticas utilizadas, os procedimentos de realização do trabalho e a descrição da meta-análise. Após isso, são descritos os resultados, com uma subseção para cada condicionamento social. Por fim, apresentam-se as considerações finais, reportando as contribuições e as limitações do presente estudo para a Sociolinguística Variacionista.

1. A NEGAÇÃO NO PORTUGUÊS BRASILEIRO

A negação é um fenômeno recorrente nas línguas humanas, ocorrendo no português brasileiro principalmente por meio da partícula *não* pré-verbal, como em (2) – exemplo criado por introspecção. Tratada muitas vezes como circunstância adverbial, *não* é um elemento que pode até

substituir uma oração inteira mediante um processo de antitaxe (substituição), como em (3), retirado de Bechara (2019, p. 314), de vertente normativa.

- (2) Eu não gosto de jogar bola.
(3) Ela fez os exercícios? – Não.

Tais regras crescem-se às descritas por Perini (2016), de vertente descritiva, o qual reconhece que o *não* pode vir antecedido e posposto ao verbo em um processo de Dupla Negação, no qual há concordância negativa (*não se anulam*). O autor reconhece que o segundo *não* ocorre geralmente no final do período, como (4), e, quando há processos de subordinação, esse *não* final nega toda a sentença, e não apenas a oração subordinada, como (5), exceto em casos de coordenação, se neles o segundo *não* vier posposto ao verbo da primeira oração, como (6)².

- (4) Eu não vou lá *não*.
(5) O Ferreira *não* sai de casa quando está chovendo *não*.
(6) Você *não* vai pegar o carro *não* e seu pai vai saber disso.

Segundo Castilho (2019, p. 577), a Negação Pós-verbal surge do desgaste fonológico do *não* anterior ao verbo em *num* até a sua cliticização em *n'* e posterior apagamento. Esse processo é teoricamente descrito como Ciclo de Jespersen, em que se hipotetiza que as línguas humanas passam por um processo no qual a Negação Pré-verbal sofre redução fonológica e, por isso, há o redobramento da negação para, por fim, a Negação Pré-verbal ser apagada em favor da Pós-verbal, que é foneticamente mais forte.

Dentro dos estudos linguísticos formalistas de orientação gerativista, a negação ocupa um espaço privilegiado. Mito, Silva e Lopes (2018, p. 63-64) argumentam que o *não* ocupa, na estrutura argumental, um sintagma próprio (NegP), que se situa, na árvore sintática, entre AgrP (Sintagma de Concordância) e TP (Sintagma de Tempo). Assim como os outros sintagmas, NegP domina o sintagma verbal, gravitando em torno dele. Os autores também comentam sobre a natureza funcional dessa categoria, que seleciona categorialmente outros elementos como complementos ou especificadores, impossibilitados de fazer seleção semântica, própria dos sintagmas lexicais.

Cavalcante (2007) analisa as três estruturas negativas – [NEG V], [V NEG] e [NEG V NEG] – e propõe derivações sintáticas para cada uma. Em relação às suas diferenças, o autor sustenta que o

2 Exemplos retirados de Perini (2016, p. 171-172).

não pré-verbal funciona como núcleo, enquanto o *não* pós-verbal, sempre tônico, é uma categoria máxima XP que ocupa a posição de especificador. Outra dessemelhança se dá no nível discursivo, em que [NEG V] é uma estrutura neutra e [NEG V NEG] e [V NEG] são estruturas pressuposicionais, que exigem material discursivo anterior inferível.

A necessidade desse material discursivo encontra base também em pesquisas sociolinguísticas sobre a negação verbal. Yacovenko e Nascimento (2016) analisaram uma amostra de Vitória (ES) e constataram que a Negação Pós-verbal é favorecida com ativação direta de proposições anteriores, como resposta a perguntas. Santos, Araújo e Pereira (2018) também constataram que a Negação Pós-verbal é favorecida em contextos de perguntas e respostas, o que pode se dar pela natureza tônica desses contextos, favorecendo a ênfase no *não*.

De maneira similar, Schwenter (2005) reporta uma base pragmática para a negação. Segundo o autor, embora haja o mesmo sentido referencial evocado pelas três variantes, elas não ocorrem nos mesmos contextos pragmáticos. Schwenter (2005, p. 1435-1436) demonstra que a Dupla Negação, ao contrário da Negação Pré-verbal, precisa da ativação de uma informação contextual – dita³, como (7), ou inferível⁴, como (8) – para ocorrer, ou seja, não é enunciada quando a sentença é uma informação nova, seja na fala ou no pensamento do interlocutor. Já a Negação Pós-verbal, por outro lado, seria mais restrita do que a Dupla Negação, exigindo que a informação contextual seja explicitamente dita, como (8). Segundo Schwenter (2005, p. 1450), a restrição aparenta ser bastante rígida, de modo a ocorrer com a repetição do verbo da pergunta.

(7) Você gostou da festa da Ana?⁵
(Não) gostei não.

(8) Você gostou da festa da Ana?
Não fui não.

Esse modelo, entretanto, não é incontroverso. Dados de Rocha (2012) apontam que, na variedade paulistana, ocorrem casos nos quais a Negação Pós-verbal é proferida quando a informação contextual é inferível. Em Freitag e Pinheiro (2020), a variável “estatuto informacional” foi controlada para análise de dados de Sergipe, mas sem significância, o que sugere não ser um fator tão influente na variação morfossintática da negação verbal nessa região.

Assim, há uma complexidade da negação verbal no português brasileiro, demandando uma gama de perspectivas teóricas para dar conta de sua análise. Nesta pesquisa, ao considerar os

3 Já ocorrido no texto anteriormente (Cunha, 2023, p. 166).

4 Identificável mediante um processo de inferência (Cunha, 2023, p. 166).

5 Os exemplos (13) e (14) foram construídos introspectivamente tomando como base os exemplos de Schwenter (2005).

estudos sociolinguísticos, foi procedida uma meta-análise para desvendar como é usada cada variante quantitativamente, além de identificar quais restrições sociais são influentes no condicionamento da negação.

2. FUNDAMENTAÇÃO TEÓRICA

O presente trabalho parte da Sociolinguística Variacionista (Labov, 2008), que propõe investigar a língua em seu contexto de uso. Nessa linha, entende-se que a língua não apresenta apenas regras categóricas – regras que se aplicam sempre da mesma maneira –, mas também regras variáveis, que são as formas que podem se alternar a depender de determinadas restrições (extra)linguísticas. Destarte, a Sociolinguística Variacionista se debruça sobre a heterogeneidade ordenada da língua ao observar que as variedades não são caóticas, mas apresentam regularidades passíveis de estudo científico (Mollica; Braga, 2020).

Justamente por serem regulares, há algo guiando a variação entre as formas linguísticas variantes. Para descrever como essas formas se comportam na língua, o “problema da restrição” focaliza na descrição de como fatores linguísticos e sociais se inter-relacionam para influenciar o falante a fazer uso de determinada forma (Labov, 2008). Nesta pesquisa, por exemplo, a negação verbal é um fenômeno variável e as suas variantes, “negação pré-verbal”, “dupla negação” e “negação pós-verbal”, são colocadas em uso em diferentes contextos, estes modulados por restrições sociais e linguísticas. São essas restrições sociais que são investigadas na meta-análise a fim de descobrir se elas realmente *restringem* o uso da negação verbal.

O fenômeno linguístico variável é codificado, na análise, como sendo uma variável dependente (VD) ternária, pois há três formas de negação verbal possíveis no PB. A partir da VD, definem-se as possíveis restrições, entendidas como variáveis independentes (VI). Esta meta-análise, em especial, foca apenas na relevância das VIs sociais como mecanismos de restrição: sexo/gênero, escolaridade e faixa etária; as quais são amplamente utilizadas em estratificação de amostras, devido ao modo como a pesquisa em Sociolinguística no Brasil se desenvolveu.

A variável sexo/gênero é frequentemente utilizada nos estudos sociolinguísticos para investigar as pressões sociais na variação. Socialmente, homens e mulheres podem ocupar espaços diferentes na sociedade, participar de grupos diferentes e, portanto, apresentar comportamentos distintos. Conforme Labov (2008, p. 348), “[...] a diferenciação sexual da fala frequentemente desempenha um papel importante no mecanismo da evolução linguística”.

Em diversos fenômenos do português, sexo/gênero é um fator estatisticamente significativo para explicar a variação de formas alternantes. Um exemplo disso é a pesquisa de Scherre (1996), que analisa a concordância nominal e reporta que mulheres utilizam mais a forma prestigiada do que homens, levando à conclusão de que as mulheres preferem formas mais conservadoras. De forma

contrária, Paiva (2020) mostra que um estudo de William Labov (1966) com o inglês de Nova Iorque mostrou que mulheres lideram processos de mudança linguística ao utilizarem a forma inovadora da pronúncia retroflexa do [r] pós-vocálico, variante prestigiada na comunidade. Nesse sentido, não parece haver concordância entre alguns dos estudos sobre o papel dos sexos/gêneros na variação e na mudança linguísticas – até porque sociedades diversas têm comportamentos diversos.

Entretanto, nem todos os fenômenos sofrem restrições impostas pelo sexo/gênero do falante. Variáveis morfossintáticas parecem sofrer mais pressões intralinguísticas do que extralinguísticas de um modo geral (cf. Silva, 2020, p. 71). Como exemplo, há a revisão baseada em meta-análise realizada por Novais e Siqueira (2020) em seis trabalhos sociolinguísticos realizados com uma amostra coletada na região do alto sertão alagoano (Vitório, 2020). O artigo reporta que apenas um trabalho apresentou significância do fator sexo/gênero, e mesmo assim com um tamanho de efeito muito baixo (V^2 de Cramer = 0,08), o que indica haver baixa associação entre a variável dependente e a independente. Logo, conforme Paiva (2020) pontua, as transformações na sociedade podem levar a uma neutralização ou acentuação da variável sexo/gênero.

Já a variável faixa etária é usada para demonstrar se há ou não um processo de mudança linguística ocorrendo, pois, como, em seu estudo de tempo aparente, é estratificada em faixas que englobam jovens, adultos e idosos, ela possibilita visualizar como falantes de diferentes gerações utilizam determinada forma. Conforme Naro (2020, p. 43), “nos eixos sociais, por exemplo, os falantes mais velhos costumam preservar mais as formas antigas [...]”, o que mostra uma diferenciação etária dentro da comunidade, pois a fala de uma pessoa reflete o seu processo de aquisição da linguagem, em uma das hipóteses sobre a mudança linguística. Esse tipo de análise dos processos de mudança é chamado de Mudança em Tempo Aparente (Naro, 2020, p. 45).

No caso da variável escolaridade, sua influência pode revelar a estigmatização de determinadas formas linguísticas pelo ambiente escolar. Especificamente no Brasil, a abordagem normativa e excludente é muito influente no ensino, em um processo no qual mudanças que favoreçam uma abordagem menos prescritivista são vistas como tentativas de empobrecimento da língua, como foi visto na polêmica acerca do livro didático “Por uma vida melhor” (Cavalcanti, 2013).

Com isso, a educação formal tende a diminuir o uso do vernáculo dos estudantes que lá chegam em prol de aproximá-los da norma-padrão, vista como o “modo certo” de se comunicar. Assim, os alunos veem sua variedade como errado e tendem a rejeitá-la. Vários exemplos disso podem ser notados a partir dos traços descontínuos, ou seja, aspectos linguísticos variáveis estigmatizados socialmente e que não perpassam todas as camadas sociais (Bortoni-Ricardo, 2004, p. 53). O uso da concordância verbal não redundante, como em “nós vai”, por exemplo, é julgado como vindo de um indivíduo não escolarizado, logo tende a ser desusado pelos alunos à medida que avançam nos estudos.

Não obstante, a não estigmatização de determinadas variantes pode levar a uma neutralização da variável escolarização, já que a escola não a tratará como forma errada de se falar o português.

Assim, os estudos com essa variável que investigam traços graduais, isto é, formas que perpassam todas as camadas sociais, tendem a encontrar pouca ou nenhuma significância estatística.

A partir do exposto, então, é necessário investigar como as três restrições sociais interferem na variação morfossintática com uso de ferramentas de comparação de trabalhos. Na seção seguinte, detalha-se como isso é feito, enfatizando o poder generalizador do método empregado: a meta-análise.

3. A META-ANÁLISE

De modo a investigar estatisticamente a relevância das VIs sexo/gênero, escolaridade e faixa etária, foi usada a meta-análise (Freitag, 2021), uma técnica de análise de revisões que agrupa dados das pesquisas e aplica a eles testes estatísticos homogêneos a fim de medir o poder explanatório geral, além de identificar correlações ou contradições entre as pesquisas. Conforme Freitag (2021), a utilização desse método auxilia na eliminação de vieses de revisões narrativas, que analisam os trabalhos individualmente. Ademais, segundo Lovatto *et al.* (2007, p. 286), “a meta-análise permite, em caso de resultados aparentemente discordantes, obter uma visão geral da situação”, o que justifica o uso desse método.

Acerca do passo a passo de uma pesquisa que utiliza a meta-análise, Lovatto *et al.* (2007) apontam cinco etapas a serem seguidas: definição do objetivo, sistematização das informações, codificação de dados, filtragem dos dados e análise dos dados. Conforme os autores, para se delimitar o objetivo da pesquisa, deve-se analisar as variáveis dependentes e independentes pesquisadas, relacionando causas e efeitos de modo a se entender o fenômeno. A partir daí, faz-se a busca sistematizada dos trabalhos, controlando o tempo (período de publicação dos textos) e o espaço (região estudada), de modo a se ter uma quantidade de trabalhos alinhada aos objetivos e próxima da realidade atual. Em seguida, deve-se organizar os trabalhos encontrados e filtrá-los a partir dos objetivos delimitados através de uma leitura crítica para encontrar eventuais problemas nas informações reportadas. Por fim, com os materiais organizados e filtrados, passa-se à análise estatística destes. Assim, percebe-se o caráter sistemático da meta-análise, que, de acordo com Lovatto *et al.* (2007, p. 293), “[...] é superior às formas tradicionais de revisão de literatura” por ser mais precisa.

Entretanto, mesmo sendo comum em outras ciências, na Linguística a meta-análise ainda é um método de recente uso, com poucos trabalhos sobre o assunto. Novais e Siqueira (2020) analisam a relevância da variável sexo/gênero em trabalhos alagoanos que partem de um mesmo *corpus*. Além dele, Siqueira (2021) investigou, por meio de meta-análise, estudos sobre o sintagma nominal possessivizado em trabalhos que utilizam o *corpus* do NURC junto a dados de um *corpus* sergipano. Outro trabalho nessa perspectiva é o estudo sobre a concordância verbal e sua relação com o contínuo rural-urbano realizado por Araújo e Freitag (2021). Outra análise que emprega esse

método de revisão é trabalho de Vitório (2024), que analisa a difusão de *a gente* na posição de sujeito na variedade alagoana, tomando por base a influência de três variáveis sociais, a saber, comunidade, escolaridade e faixa etária.

Apesar de incipiente, esse método de pesquisa mostra-se frutífero por agrupar os diversos estudos sobre os fenômenos variáveis e proporcionar generalizações mais confiáveis acerca da variação e mudança linguística. Nesse sentido, a meta-análise é uma ferramenta proveitosa de ser utilizada e, por isso, é o método escolhido neste trabalho.

3.1. MEDIDAS ESTATÍSTICAS

Nesta meta-análise, alguns termos e testes estatísticos foram realizados, logo, necessitam de explicação. O primeiro é o valor de p , um valor de significância que diz respeito à probabilidade de um evento ter ocorrido ao acaso, convencionalmente 5% (Oushiro, 2017, p. 83). Em outras palavras, a cada 100 análises estatísticas feitas, 5 delas teriam um resultado diferente do encontrado. Com $p > 0,05$, diz-se que a relação entre uma dada VI e uma VD não é significativa; com $p \leq 0,05$, diz-se que é significativa (há uma possibilidade minúscula de o resultado ser ao acaso).

Entretanto, o valor de p é sensível ao tamanho amostral e ao teste empregado, por isso, pesquisas que tendem a comparar a significância de variáveis podem ter problemas se as amostras forem de tamanhos distintos. Atenuar esse problema implica utilizar uma medida conhecida como tamanho de efeito, que analisa de modo padronizado a magnitude de um determinado fenômeno (Lee, 2016). No caso desta pesquisa, um dos tamanhos de efeito utilizados é o V^2 de Cramer, um valor de correlação de 0 a 1 calculado com base no teste de qui-quadrado (teste de associação entre duas variáveis). Lee (2016) dispõe uma tabela de interpretação do valor de V^2 de Cramer que varia de “Insignificante” (0,0–0,1) a “Muito forte” (0,8–1,0). Outro valor de tamanho de efeito é o R^2 , que também varia de 0 a 1 e é utilizado na análise dos modelos de regressão na seção de resultados.

Junto ao tamanho de efeito, a Estatística dispõe de uma medida para generalização dos resultados chamada de Intervalo de Confiança. Ela é um meio de generalizar os resultados para a população com 95% de certeza (Gries, 2019, p. 127). O intervalo de confiança estima um valor de limite inferior e um de limite superior para indicar que, entre 100 intervalos de confiança calculados, o valor populacional seria encontrado em 95% deles (Lee, 2016). Quanto mais próximo é o valor dos limites inferior e superior, mais significativo o resultado encontrado; quanto mais distante, menos significativo.

Outro método utilizado nesta pesquisa é a regressão logística. A regressão logística é um modelo em que o logaritmo das razões das chances (a divisão dos “sucessos” pelos “fracassos”), ou *log-odds*, é estimado pela alteração do nível de uma variável em comparação a outra – quando a variável dependente possui mais de dois níveis, a regressão utilizada é a multinomial, que é um tipo de regressão logística. Nesta meta-análise, se a regressão logística tiver como nível de referência a

Negação Pré-verbal, o modelo estima o quão significativa é uma determinada VI sobre a Dupla Negação, por exemplo (em comparação ao nível de referência).

Nesta pesquisa, todo o procedimento estatístico é realizado com auxílio do RStudio, uma interface do R, que é uma linguagem de programação utilizada em muitas pesquisas estatísticas pelo seu maquinário desenvolvido para análises de dados. No RStudio, foram utilizados os seguintes pacotes estatísticos: *ggstatsplot* (Patil, 2021), para criar gráficos; *sjPlot* (Lüdtke, 2023), para criação das tabelas com modelos de regressão; *tidyverse* (Wickham *et al.*, 2019), para utilizar ferramentas de filtragem de dados; e *nnet* (Venables; Ripley, 2002), para a feitura das regressões multinomiais.

4. REVISÃO SISTEMÁTICA DE LITERATURA

Realizou-se uma busca avançada no *Google Acadêmico*, nos meses de maio e junho de 2022, em que se procurou, nos títulos⁶, textos que tivessem a palavra “negação” e que pudessem ter pelo menos uma das seguintes palavras: Português, português⁷, Brasil, Nordeste, Sudeste, pré-verbal, dupla e pós-verbal. Na busca, não foi definido período de tempo e foram consultadas todas as páginas de resultado. Com esse procedimento, foram encontrados 149 resultados (com o filtro “incluir citações” desativado para que não houvesse resultados de citações de trabalhos).

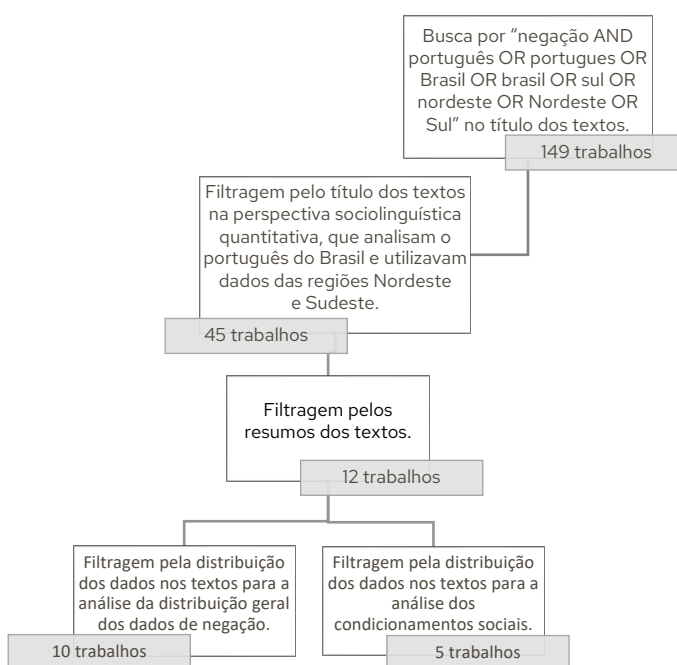
A partir daí, passou-se à análise do título dos materiais em busca de filtrar os que se encaixam na perspectiva sociolinguística quantitativa, analisam o português do Brasil e utilizavam dados das regiões Nordeste e Sudeste. Como resultado, 45 trabalhos foram filtrados. Após isso, procedeu-se uma análise dos resumos dos textos para filtrar ainda mais os trabalhos com os mesmos critérios anteriores, além de se acrescentar o critério de o arquivo conter o texto inteiro, e não apenas uma parte. Com isso, foram selecionados 12 trabalhos que se encaixavam em uma investigação quantitativa e sociolinguística da variação dos três tipos de negação verbal presentes no português brasileiro, com base em dados das regiões Nordeste e Sudeste. Foram eles: Braga e Silva (2011); Castro *et al.* (2017), Cavalcante (2007), Lopes da Silva (2020), Nascimento (2014), Nunes (2014), Reimann e Yacovenco (2011, 2014), Rocha (2012, 2013), Serra (2018) e Yacovenco e Nascimento (2016).

6 O Google Acadêmico não é a plataforma mais recomendada para revisões sistemáticas por apresentar problemas em encontrar artigos e não lidar bem com as palavras-chave e os operadores de busca (Giustini; Boulos, 2013), além de não mostrar, na maioria dos casos, nota de retratação para artigos problemáticos (Pastor-Ramon *et al.*, 2022). Entretanto, ele mostra-se relativamente adequado em pesquisas por títulos, que foi a escolhida neste trabalho, e é a base de dados disponível a pesquisadores das áreas de Humanas, visto que bases mais eficientes, como o PubMed, são voltadas à área da Saúde. Nesse sentido, o Google Acadêmico, mesmo com seus problemas, ainda é relevante para a realização de revisões sistemáticas dentro da Linguística.

7 Alguns trabalhos fazem uso da palavra com inicial minúscula, então foram colocadas as duas formas.

O passo final da filtragem de trabalhos foi analisar quais disponibilizam os dados completos (isto é, as contagens) para seu uso na meta-análise, além de conterem amostras de dados que sejam significativamente numerosas (não contenham dados de apenas um ou dois informantes⁸). Com isso, Braga e Silva (2011) e Castro *et al.* (2017) foram descartados por terem apenas dois informantes. Dos 10 trabalhos que restaram, todos foram utilizados para uma análise geral da distribuição da negação. Entretanto, Reimman e Yacovenko (2011) foi descartado na análise dos condicionamentos sociais por amalgamar os dados da Dupla Negação com a Negação Pós-verbal; também Reimann e Yacovenko (2014), Rocha (2012), Serra (2018) e Yacovenko e Nascimento (2016) foram descartados nessa análise por não apresentarem os dados das variáveis sociais (não foram tidas como significativas ao serem rodadas no GoldVarb X, o que já revela o condicionamento nulo dessas variáveis). Com isso, os trabalhos que foram utilizados na meta-análise dos condicionamentos sociais são: Cavalcante (2007), Lopes da Silva (2020), Nascimento (2014), Nunes (2014) e Rocha (2013).

O fluxograma a seguir indica graficamente como foi feita a filtragem dos trabalhos.



Fluxograma 1. Revisão sistemática de literatura. Fonte: Santos (2023).

As pesquisas selecionadas utilizam *corpora* diferentes para fazer a análise da negação verbal. Lopes da Silva (2020) utiliza um *corpus* coletado por outra pesquisadora de entrevistas

⁸Para evitar uma diferença significativa de informantes entre os trabalhos.

sociolinguísticas semiestruturadas com informantes de Fortaleza (CE). Nunes (2014) utiliza uma subamostra do Programa de Estudos sobre o Uso da Língua (PEUL), que descreve o comportamento linguístico no Rio de Janeiro. Reimann e Yacovenço (2011, 2014), Nascimento (2014) e Yacovenço e Nascimento (2016) utilizam amostras do projeto Português Falado na Cidade de Vitória (PortVix), composto de 46 entrevistas de informantes nascidos em Vitória (ES). Rocha (2012, 2013) utiliza amostra de 48 entrevistas coletadas de modo semialeatório de informantes nascidos na cidade de São Paulo (*corpus* GESOL-USP). Cavalcante (2007) utiliza uma amostra do *corpus* Projeto Vertentes do Português Rural da Bahia, que coleta dados de comunidades afrodescendentes do interior da Bahia. Serra (2018) coletou um *corpus* próprio de 32 informantes (dos quais 24 foram usados na pesquisa) referente a falantes nativos de São Luís e Jamary dos Pretos (MA). Todos eles são analisados na seção seguinte de Análise Geral.

5. ANÁLISE GERAL

Das 10 pesquisas selecionadas para análise geral a fim de entender o comportamento da negação verbal nas regiões Nordeste e Sudeste, a tabela 1 mostra a distribuição da negação em cada uma delas. Em todas as regiões, nota-se uma prevalência da Negação Pré-verbal seguida da Dupla Negação e, por fim, da Negação Pós-verbal. Esta, por sua vez, tem atuação limitada, provavelmente devido aos seus contextos de uso descritos por Schwenter (2005). Pela tabela, também é possível perceber maior número de Negação Pós-verbal ($n = 115$, 5,6%) na pesquisa de Cavalcante (2007), que foi feita em uma comunidade afrodescendente baiana.

| Pesquisa | Região | Estado | Amostra | Infor- mantes | Pré- verbal | Dupla Negação | Negação Pós-verbal | Total |
|-------------------------------|----------|----------------|---------------|------------------|----------------|------------------|-----------------------|-------------|
| Lopes da Silva (2020) | Nordeste | Ceará | Torres (2008) | 40 | 1819 (86,7%) | 278 (13,2%) | 7 (0,3%) | 2097 (100%) |
| Nunes (2014) | Sudeste | Rio de Janeiro | PEUL | 8 | 616 (73%) | 214 (25,3%) | 13 (1,5%) | 843 (100%) |
| Nascimento (2014) | Sudeste | Espírito Santo | PortVix | 18 | 1751 (77,3%) | 478 (21,1%) | 34 (1,5%) | 2263 (100%) |
| Rocha (2013) | Sudeste | São Paulo | GESOL | 48 | 5279 (94,1%) | 324 (5,7%) | 4 (0,07%) | 5607 (100%) |
| Cavalcante (2007) | Nordeste | Bahia | Vertentes | 18 | 1343 (66,2%) | 568 (28%) | 115 (5,6%) | 2026 (100%) |
| Serra (2018) | Nordeste | Maranhão | Própria | 24 | 1009 (87,5%) | 133 (11,4%) | 17 (1,4%) | 1159 (100%) |
| Reimann e Yacovenço (2011) | Sudeste | Espírito Santo | PortVix | 8 | 721 (73,6%) | 216 (22%) | 42 (4,2%) | 979 (100%) |
| Rocha (2012) | Sudeste | São Paulo | GESOL | 12 | 940 (88,6%) | 117 (11%) | 4 (0,4%) | 1061 (100%) |
| Yacovenço e Nascimento (2016) | Sudeste | Espírito Santo | PortVix | 18 | 1751 (77,3%) | 478 (21,1%) | 34 (1,5%) | 2263 (100%) |
| Reimann e Yacovenço (2014) | Sudeste | Espírito Santo | PortVix | 18 | 1754 (77,4%) | 478 (21,1%) | 34 (1,5%) | 2266 (100%) |
| Lopes da Silva (2020) | Nordeste | Ceará | Torres (2008) | 40 | 1819 (86,7%) | 278 (13,2%) | 7 (0,3%) | 2097 (100%) |

Tabela 1. Distribuição da negação nas pesquisas selecionadas pela revisão sistemática Fonte: Santos (2023).

Outro aspecto a se notar nessa comparação é a semelhança entre as pesquisas de Yacovenko e Nascimento (2016) e Reimann e Yacovenko (2014), com uma diferença de apenas 3 dados na Negação Pré-verbal. Nesse sentido, percebe-se que as autoras publicaram a pesquisa, inicialmente, em Anais (Reimann; Yacovenko, 2014) e, depois, publicaram em um periódico (Yacovenko; Nascimento, 2016). Para visualização e teste de hipótese, o gráfico 1 apresenta os resultados sem considerar Yacovenko e Nascimento (2016).

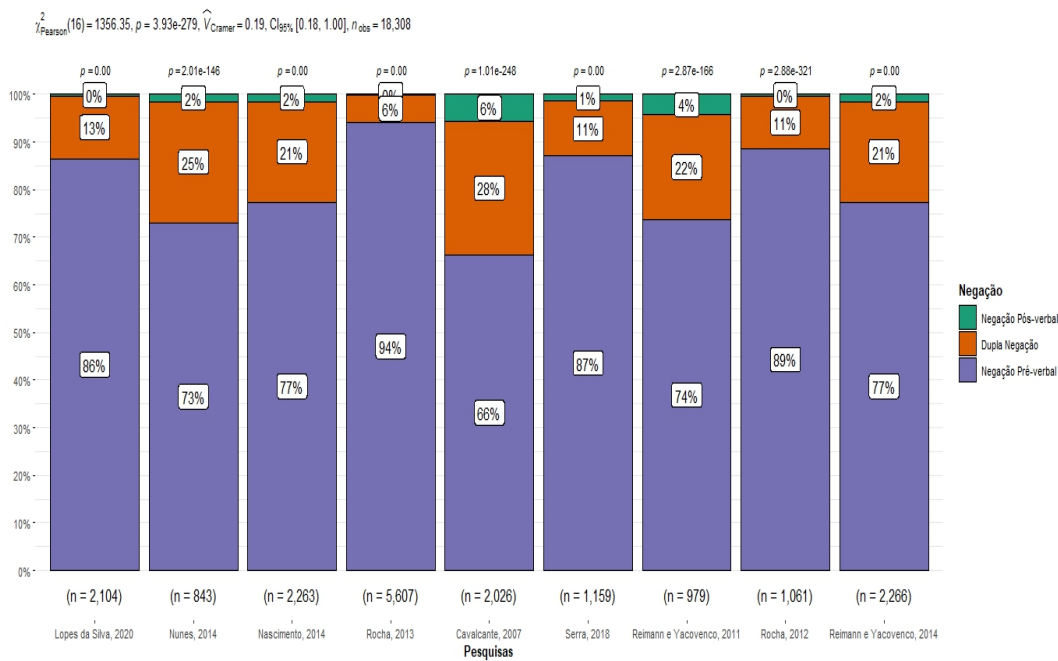


Gráfico 1. Distribuição das negações por pesquisa selecionada na revisão sistemática. Fonte: Santos (2023).

A diferença entre os resultados dos trabalhos é significativa ($p < 0,05$). Embora a Negação Pré-verbal seja a mais usada, o percentual varia em cada pesquisa. As que menos apresentam dados são Cavalcante (2007), no Nordeste, e Nunes (2014), no Sudeste. Da mesma maneira, são estas as duas pesquisas que mais apresentam dados para a Dupla Negação. Sobre a Negação Pós-verbal, Cavalcante (2007) e Reimann e Yacovenko (2011) – também do Sudeste – apresentam mais dados. Nesse sentido, não se nota uma convergência de resultados para uma ou outra região, visto que Lopes da Silva (2020), no Ceará, e Serra (2018), no Maranhão, apresentam baixíssimo percentual para essa variante. Assim, a hipótese de a Negação Pós-verbal e Dupla Negação serem mais favorecidas no Nordeste em razão do contato linguístico não se sustenta.

5.1. ANÁLISE DA VARIÁVEL SEXO/GÊNERO

Na análise da variável sexo/gênero, o procedimento inicial seria agregar, em um gráfico comparativo, todas as pesquisas que falavam da negação verbal, já que todas tinham a mesma delimitação (masculino, feminino). Entretanto, a variável dependente era analisada de modo ternário em três trabalhos, enquanto dois deles analisavam a variação entre a Negação Pré-verbal e a Dupla Negação (porque não encontraram dados suficientes para a análise ser feita com o terceiro tipo de negação).

Para resolver esse problema, adotou-se outra estratégia. Como a análise multinomial divide os dados em duas regressões logísticas juntas, foi feita a separação para serem duas as análises: a primeira delas analisando a variação entre a Negação Pré-verbal e a Dupla, que foi feita com os cinco trabalhos, e a segunda análise focando na variação entre a Negação Pós-verbal e a Dupla Negação⁹, que utilizou três trabalhos. Assim, foi possível analisar como cada negação variava em relação ao nível de referência “Negação Pré-verbal”, usando como modelo para fazer a regressão logística a função “glm (VD ~ genero, data=x, family = 'binomial')”, sendo x o conjunto de dados de cada pesquisa¹⁰.

A influência da variável “sexo/gênero” no condicionamento da Dupla Negação está exposta no gráfico 2. Nele, foi utilizado como valor de intercepto (β) o quanto variava em *log-odds* pela alteração do valor de referência “masculino” para “feminino”, pois se tem uma ideia da influência de tal variável social no uso da negação (o nível de referência é “Negação Pré-verbal”).

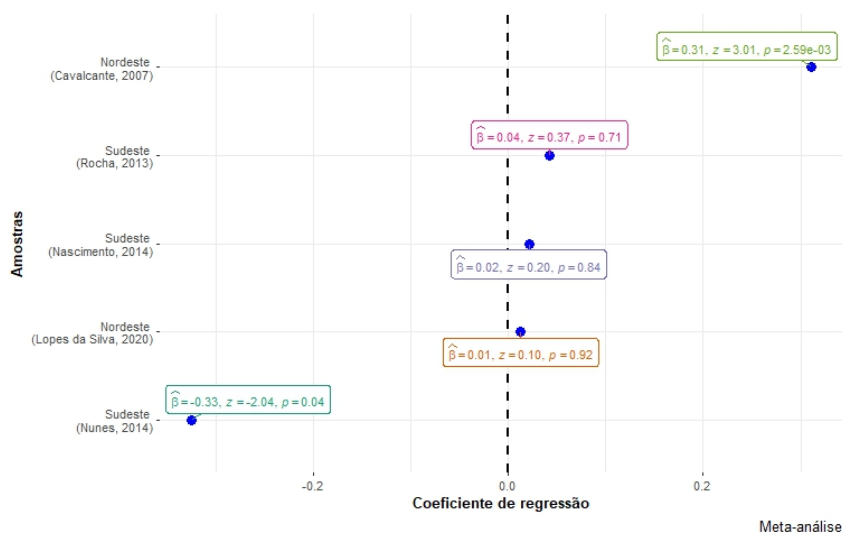


Gráfico 2. Influência da variável “sexo/gênero” na variação entre Negação Pré-verbal e Dupla Negação. Fonte: Santos (2023).

9 Foi escolhida essa relação, e não a entre Negação Pré-verbal e Negação Pós-verbal, porque as duas são utilizadas em contextos diferentes, conforme apontou Schwenter (2005).

10 O *script* completo pode ser consultado no GitHub, cujo endereço de acesso está na seção Declaração de Disponibilidade de Dados.

Como se vê, a variável sexo/gênero só foi significativa nos trabalhos de Nunes (2014) e Cavalcante (2007). Se se tomasse como parâmetro a análise desses dois estudos apenas, concluir-se-ia que a variável sexo/gênero é influente – já que indicariam que, no Nordeste, as mulheres favoreceriam a Dupla Negação e, no Sudeste, haveria favorecimento da Negação Pré-verbal. Entretanto, ao inserir mais trabalhos sobre o tema, percebe-se que a maioria deles não trouxe efeito significativo – mesmo Nunes (2014) tem um valor de p muito próximo do limiar convencional de 5%.

Além disso, outro fator que pode influenciar na significância da variável em Cavalcante (2007) é a própria comunidade analisada. Nesse trabalho, analisou-se a negação com base em um *corpus* retirado do Projeto Vertentes do Português Rural da Bahia (Cinzento, Rio das Contas e Sapé), em que a maioria da população é de ascendência negra. Conforme o autor, os homens acabam se deslocando mais para o centro, enquanto as mulheres ficam mais em casa, o que favorece a divisão sexual do trabalho (Cavalcante, 2007, p. 68). Nesse sentido, a alteração de papéis sociais nessa região rural acaba aproximando mais os homens do padrão convencional na língua, diferenciando os usos linguísticos entre os sexos.

Já a influência da variável “sexo/gênero” sobre a variação entre a Dupla Negação e a Negação Pós-verbal foi analisada no gráfico 3. Da mesma maneira, utilizou-se como β o quanto variava em *log-odds* pela alteração do nível da variável (sendo o nível de referência “Dupla Negação”).

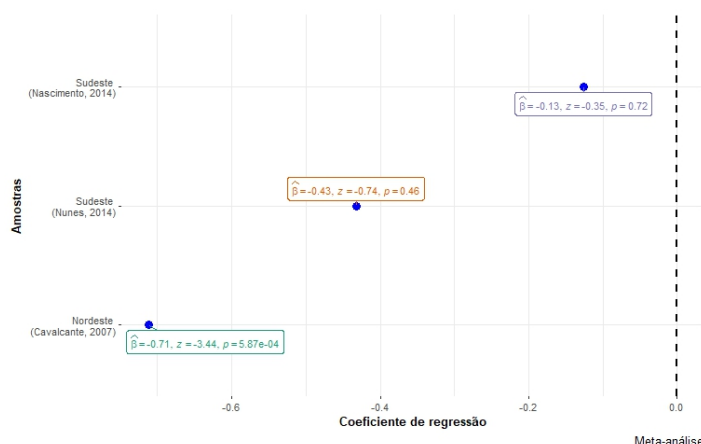


Gráfico 3. Influência da variável “sexo/gênero” na variação entre a Dupla Negação e a Negação Pós-verbal. Fonte: Santos (2023).

Nos três trabalhos que apresentavam dados para a Negação Pós-verbal em cada sexo/gênero, nota-se a mesma realidade. O único trabalho com resultado significativo dentro do modelo é o de Cavalcante (2007), que encontrou um número alto de dados de Negação Pós-verbal. Nos resultados dos outros trabalhos, entretanto, o efeito da variável foi baixo ou nulo. Devido a isso, é cabível pensar que novos *corpora* desenvolvidos para analisar a negação verbal poderiam buscar outras alternativas para a segmentação das células sociais além de/em vez de sexo/gênero, visto

que, conforme as análises da presente meta-análise, independentemente de ser homem ou mulher, o informante tenderia a utilizar a negação de maneiras similares, seja no Nordeste ou no Sudeste.

5.2. ANÁLISE DA VARIÁVEL ESCOLARIDADE

Na análise da variável escolaridade, comparar os trabalhos em um único gráfico não foi possível porque todos eles utilizavam diferentes níveis. Devido a isso, os estudos são analisados individualmente e, no fim, faz-se uma comparação geral. As regressões logísticas e multinomiais foram feitas com os modelos estatísticos “glm (VD ~ escolaridade, data= x, family = 'binomial')” e “multinom (VD ~ escolaridade, data= x)”, respectivamente, sendo x o conjunto de dados de cada pesquisa¹¹.

A primeira pesquisa analisada foi a de Cavalcante (2007). Nesse trabalho, o autor investigou dois níveis de escolaridade: analfabetos e semialfabetizados (pessoas que tiveram algum tipo de contato com a escolarização). A tabela 2 apresenta os resultados, que foram significativos (o nível de referência é “Semialfabetizados”):

| <i>Predictors</i> | <i>Log-Odds</i> | <i>CI</i> | VD | | <i>Response</i> |
|--|-----------------|---------------|-----------|--|--------------------|
| | | | <i>p</i> | | |
| (Intercept) | -1.07 | -1.20 – -0.94 | <0.001 | | Dupla Negação |
| escolaridade [Analfabetos] | 0.53 | 0.33 – 0.73 | <0.001 | | Dupla Negação |
| (Intercept) | -2.63 | -2.88 – -2.37 | <0.001 | | Negação Pós-verbal |
| escolaridade [Analfabetos] | 0.41 | 0.02 – 0.79 | 0.040 | | Negação Pós-verbal |
| Observations | 2026 | | | | |
| R ² / R ² adjusted | 0.009 / 0.008 | | | | |

Tabela 2. Influência da variável “escolaridade” em Cavalcante (2007). Fonte: Santos (2023).

Curiosamente, essa variável não foi significativa na análise do autor, o que pode se dar por sua pesquisa envolver uma regressão logística multivariada, enquanto aqui é analisada individualmente. Percebe-se que menor nível de escolarização aparenta estar relacionado a um favorecimento de 0,53 *log-odds* no uso da Dupla Negação quando comparada à Negação Pré-verbal. Da mesma forma, o uso da Negação Pós-verbal é favorecido em 0,41 *log-odds*, mas é preciso estar atento ao intervalo de confiança, cujo limite inferior aproxima-se muito do zero.

Em nível de probabilidade, os resultados ficam mais evidentes. O uso da Dupla Negação quando a escolaridade do falante é “semialfabetizado” é de -1,07 *log-odds*, o que equivale a uma

¹¹ Mais detalhes sobre o código podem ser consultados no GitHub.

probabilidade de 25,5% de uma forma dessa aparecer no discurso de alguém com mais escolarização¹². Já, em analfabetos, os *log-odds* de uso da Dupla Negação são de -0,54, o que equivale a uma probabilidade de 36,8% de ocorrência dessa forma linguística. Com a Negação Pós-verbal, os *log-odds* iniciais de -2,63 equivalem a 6,7% de chance de ocorrência em sujeitos semialfabetizados. Em analfabetos, são -2,22 *log-odds*, ou probabilidade de 9,7%. Em conclusão, o uso da Dupla Negação é significativamente diferente entre os grupos, assim como o da Negação Pós-verbal – porém, em menor escala, o que influencia o valor baixo da medida de tamanho de efeito R^2 .

Em Lopes da Silva (2020), a variável escolaridade foi dividida em “ensino médio” e “ensino superior”. Os resultados dessa análise, portanto, não podem ser diretamente comparáveis aos de Cavalcante (2007), visto que este trata de um grupo de pessoas diferente. De todo modo, os resultados do modelo de regressão logística estão dispostos na tabela 3 (o nível de referência é “Ensino Médio”).

| <i>Predictors</i> | <i>Log-Odds</i> | VD | | <i>p</i> |
|--------------------------------|-----------------|---------------|--|------------------|
| | | <i>CI</i> | | |
| (Intercept) | -1.83 | -2.02 – -1.66 | | <0.001 |
| escolaridade [Ensino Superior] | -0.08 | -0.34 – 0.17 | | 0.517 |
| Observations | 2097 | | | |
| R^2 Tjur | 0.000 | | | |

Tabela 3. Influência da variável “escolaridade” em Lopes da Silva (2020). Fonte: Santos (2023).

Nessa pesquisa, como foi utilizada uma variável dependente binária, a regressão produzida foi a logística. Na tabela 3, percebe-se que a escolaridade não é influente, ou seja, falantes do ensino médio e do ensino superior usam a negação do mesmo modo. Nesse contexto, a norma-padrão parece não exercer pressão social sobre os indivíduos, o que dá indícios indiretos de uma não estigmatização na sociedade. Ademais, outras medidas relevantes na tabela são o R^2 nulo e o intervalo de confiança ultrapassando o zero.

Conclui-se, portanto, que não há homogeneidade na região Nordeste acerca da influência da escolaridade na variação da negação verbal. Os trabalhos analisados se mostraram díspares, com significância estatística na Bahia (Cavalcante, 2007) e insignificância no Ceará (Lopes da Silva, 2020). O Sudeste, por sua vez, foi estudado em Rocha (2013), Nunes (2014) e Nascimento (2014). De início, analisam-se os dados de Rocha (2013), que estudou a comunidade da capital paulista,

12 Transformação de *log-odds* para probabilidade feita a partir da função `ilogit()`. Conferir Oushiro (2017) para mais detalhes.

utilizando como níveis da variável escolaridade “ensino médio” e “ensino superior” (o nível de referência utilizado foi “Ensino Médio”).

| <i>Predictors</i> | VD | | |
|--------------------------------|-----------------|---------------|----------|
| | <i>Log-Odds</i> | <i>CI</i> | <i>p</i> |
| (Intercept) | -2.51 | -2.66 – -2.37 | <0.001 |
| escolaridade [Ensino Superior] | -0.62 | -0.85 – -0.39 | <0.001 |
| Observations | 5603 | | |
| R ² Tjur | 0.005 | | |

Tabela 4. Influência da variável “escolaridade” em Rocha (2013). Fonte: Santos (2023).

Como se observa, a variável mostrou-se significativa. Nesse caso, a Dupla Negação mostrou-se mais desfavorecida quando a escolaridade aumentava, passando de -2,51 *log-odds* (7,1%) para -3,13 (4,41%). Ou seja, no caso da população de São Paulo (capital), pode haver uma estigmatização da Dupla Negação, mas essa hipótese carece de evidências mais robustas.

Na tabela 5, apresenta-se a regressão produzida a partir dos dados de Nunes (2014). Nesse trabalho, os níveis da VI foram “5-8 anos” e “9-11 anos” de escolarização (o nível de referência é “5-8 anos”). Nesse sentido, pode-se presumir que se trataria de sujeitos que estudaram o Ensino Fundamental e o Médio.

| <i>Predictors</i> | <i>Log-Odds</i> | <i>CI</i> | <i>p</i> | <i>Response</i> |
|--|-----------------|---------------|----------|--------------------|
| | | | | |
| (Intercept) | -0.91 | -1.11 – -0.70 | <0.001 | Dupla Negação |
| escolaridade9-11 anos | -0.35 | -0.67 – -0.03 | 0.030 | Dupla Negação |
| (Intercept) | -3.48 | -4.11 – -2.85 | <0.001 | Negação Pós-verbal |
| escolaridade9-11 anos | -1.10 | -2.40 – 0.20 | 0.098 | Negação Pós-verbal |
| Observations | 843 | | | |
| R ² / R ² adjusted | 0.007 / 0.005 | | | |

Tabela 5. Influência da variável “escolaridade” em Nunes (2014). Fonte: Santos (2023).

Observa-se na tabela que a escolarização só apresentou efeito significativo no uso da Dupla Negação, e mesmo assim com baixa significância ($p = 0,03$ e intervalo de confiança próximo ao zero). Nesse caso, os *log-odds* passam de -0,91 (28,7%) para -1,26 (22,1%). Ou seja, quanto maior a escolarização, menos tende o falante a usar a Dupla Negação no Rio de Janeiro, com base nos dados de Nunes (2014).

Por fim, a tabela 6 apresenta a regressão multinomial criada para dar conta dos dados de Nascimento (2014) (o valor de referência é “Ensino Fundamental”). Nesse estudo, a escolaridade foi

elaborada com os níveis “Ensino Fundamental”, “Ensino Médio” e “Ensino Universitário”, ou seja, abrange os três níveis de escolarização do Brasil. Nesse sentido, há uma análise mais abrangente da influência dessa variável.

| <i>Predictors</i> | <i>Log-Odds</i> | <i>CI</i> | VD | | <i>Response</i> |
|--|-----------------|---------------|-----------|--|--------------------|
| | | | <i>p</i> | | |
| (Intercept) | -1.12 | -1.29 – -0.95 | <0.001 | | Dupla Negação |
| escolaridade [Ensino Médio] | -0.31 | -0.55 – -0.08 | 0.010 | | Dupla Negação |
| escolaridade [Ensino Universitário] | -0.22 | -0.48 – 0.03 | 0.086 | | Dupla Negação |
| (Intercept) | -3.53 | -4.02 – -3.03 | <0.001 | | Negação Pós-verbal |
| escolaridade [Ensino Médio] | -1.23 | -2.17 – -0.28 | 0.011 | | Negação Pós-verbal |
| escolaridade [Ensino Universitário] | -0.23 | -0.99 – 0.53 | 0.557 | | Negação Pós-verbal |
| Observations | 2263 | | | | |
| R ² / R ² adjusted | 0.005 / 0.005 | | | | |

Tabela 6. Influência da variável “escolaridade” em Nascimento (2014). Fonte: Santos (2023).

Em Nascimento (2014), percebe-se a significância estatística apenas na alteração de escolaridade do Ensino Fundamental para o Ensino Médio. Na Dupla Negação, a tabela mostra que os *log-odds* passam de -1,12 (24,6%) para -1,43 (19,3%) quando o sujeito tem Ensino Médio – o valor de -1,12 é derivado da soma dos *log-odds* do efeito negativo do nível “Ensino Médio” ao valor do Intercepto. Na Negação Pós-verbal, por sua vez, percebe-se uma mudança de -3,53 (2,8%) para -4,76 (0,8%). Nesse sentido, percebe-se que as formas as quais apresentam a partícula negativa posterior ao verbo são desfavorecidas na comunidade de Vitória (ES).

Curiosamente, o ensino superior não se mostra um condicionamento significativo das negações que apresentam o *não* após o verbo. Em Cavalcante (2007), a escolarização foi fator significativo entre quem não havia entrado na escola e quem havia. Em Nunes (2014), a passagem de 5-8 anos (Ensino Fundamental, presume-se) para 9-11 (Ensino Médio) desfavoreceu significativamente a Dupla Negação. Na mesma linha, a variação de “ensino fundamental” para “ensino médio” foi significativa em Nascimento (2014), mas a mudança para o “ensino superior”, não – em Lopes da Silva (2020), a alteração do nível da escolaridade para “ensino superior” também não foi significativa.

O único trabalho em que houve significância da escolaridade para o ensino superior foi Rocha (2013), o que pode indicar alguma idiosincrasia de São Paulo, a princípio. Nesse sentido, é possível especular que há uma correlação entre a escolaridade e o uso da negação até o ensino fundamental, o que pode revelar a pressão social da norma-padrão sobre a língua. De todo modo, mais estudos são necessários para determinar a influência da escolarização na variação linguística, tanto da negação quanto de outros fenômenos linguísticos variáveis.

5.3. ANÁLISE DA VARIÁVEL FAIXA ETÁRIA

A variável faixa etária, de modo similar à variável escolaridade, foi estruturada de modo diferente em cada pesquisa, seja mudando os grupos por idade ou tomando duas faixas etárias em vez de três. Nesse contexto, a comparação direta com outros trabalhos fica prejudicada e opta-se pela análise individual dos trabalhos, limitando o poder explanatório da meta-análise. Para essa variável, as regressões logísticas e multinomiais foram feitas com os modelos estatísticos “glm (VD ~ faixa.etaria, data= x, family = 'binomial')” e “multinom (VD ~ faixa.etaria, data= x)”, respectivamente, sendo x o conjunto de dados de cada pesquisa¹³.

De início, há o trabalho de Cavalcante (2007) em comunidades na Bahia. Nesse estudo, o autor seccionou a faixa etária nos níveis “20-40 anos”, “41-60 anos” e “>60 anos”. Os resultados são dispostos na tabela 7 (o nível de referência é “20-40 anos”).

| Predictors | Log-Odds | CI | VD | |
|--|---------------|---------------|--------|--------------------|
| | | | p | Response |
| (Intercept) | -0.81 | -0.99 – -0.64 | <0.001 | Dupla Negação |
| faixa.etaria41-60 anos | -0.22 | -0.46 – 0.02 | 0.073 | Dupla Negação |
| faixa etaria [+60 anos] | 0.12 | -0.13 – 0.37 | 0.346 | Dupla Negação |
| (Intercept) | -2.21 | -2.52 – -1.90 | <0.001 | Negação Pós-verbal |
| faixa.etaria41-60 anos | -1.02 | -1.54 – -0.49 | <0.001 | Negação Pós-verbal |
| faixa etaria [+60 anos] | 0.13 | -0.30 – 0.56 | 0.556 | Negação Pós-verbal |
| Observations | 2026 | | | |
| R ² / R ² adjusted | 0.009 / 0.008 | | | |

Tabela 7. Influência da variável “faixa etária” em Cavalcante (2007). Fonte: Santos (2023).

Nesse trabalho, a faixa etária se mostrou significativa apenas em relação ao uso da Negação Pós-verbal, que foi desfavorecida da primeira faixa etária para a segunda, passando de -2,21 *log-odds* (9,9%) para -3,23 (3,8%). Esses resultados indicam que o fenômeno aparenta estar em estabilidade, visto que o uso da negação na faixa etária entre 20 e 40 anos é similar ao da faixa etária maior de 60 anos, com maior desfavorecimento entre a população intermediária (40-60) – o que explica o R² baixíssimo (0,009).

Já, no estudo de Lopes da Silva (2020), a faixa etária foi binária: “<30 anos” e “>30 anos”. Os resultados são expostos na tabela 8 (o nível de referência é “<30 anos”).

¹³ Mais detalhes acerca do código podem ser consultados no código disponível no GitHub.

| VD | | | |
|-------------------------|----------|---------------|--------|
| Predictors | Log-Odds | CI | p |
| (Intercept) | -2.04 | -2.23 – -1.86 | <0.001 |
| faixa etária [+30 anos] | 0.31 | 0.06 – 0.57 | 0.016 |
| Observations | 2097 | | |
| R ² Tjur | 0.003 | | |

Tabela 8. Influência da variável “faixa etária” em Lopes da Silva (2020). Fonte: Santos (2023).

Os dados da tabela mostram que uma faixa etária maior favorece o uso da Dupla Negação, que passa inicialmente de -2,04 *log-odds* (11,5%) para -1,73 (15%), um aumento pequeno, que se explica pelo intervalo de confiança próximo do zero no limite interior e pelo baixo valor do R² (0,003). Nesse sentido, embora pareça significativa, sua influência é muito baixa, e poderia ser nulificada ao se acrescentarem novas variáveis ao modelo de regressão logística criado – foi o que ocorreu na análise de Lopes da Silva (2020), a qual não encontrou significância estatística nessa variável em sua análise multivariada.

Em Rocha (2013), que analisou dados da comunidade paulista, a VI “faixa etária” foi dividida entre os níveis “18-35 anos”, “36-55 anos” e “>56 anos”. Nessa pesquisa, que também usa a VI ternária, o intervalo de anos para cada faixa etária é diferente do apresentado em Cavalcante (2007), ou seja, mais uma impossibilidade aparece para a comparação entre os trabalhos. Mesmo assim, apresentam-se os resultados da pesquisa de Rocha (2013) na tabela 9.

| VD | | | |
|-------------------------|----------|---------------|--------|
| Predictors | Log-Odds | CI | p |
| (Intercept) | -2.94 | -3.15 – -2.75 | <0.001 |
| faixa.etaria36-55 anos | 0.18 | -0.10 – 0.46 | 0.212 |
| faixa etária [>56 anos] | 0.28 | 0.01 – 0.55 | 0.046 |
| Observations | 5603 | | |
| R ² Tjur | 0.001 | | |

Tabela 9. Influência da variável “faixa etária” em Rocha (2013). Fonte: Santos (2023).

Nessa pesquisa, há um aumento do uso da Dupla Negação à medida que o falante envelhece, embora seja um valor baixo e somente significativo (bem próximo ao valor crítico de 5%) quando o falante passa dos 56 anos. Além disso, o valor de R² é praticamente nulo, assim como o intervalo de confiança, que quase toca o zero no limite inferior, indicando uma quase insignificância do resultado. Na análise multivariável de Rocha (2013), a faixa etária foi um fator com baixa significância, mostrando que o acréscimo de outras VIs tornam a faixa etária descartável. Nas palavras do autor, “o grupo *faixa etária* indica um pequeno favorecimento de NEG2 entre os informantes mais velhos

[...], bem como um pequeno desfavorecimento entre os mais jovens” (Rocha, 2013, p. 66, grifos do autor) – o uso do adjetivo “pequeno” já indica que tal VI não parece ser um condicionamento tão relevante em seu trabalho.

Nunes (2014), por sua vez, analisou a variável faixa etária pelos níveis “15-25 anos” e “26-50 anos”. De início, um ponto chama a atenção: o intervalo de anos em cada grupo é distinto; 10 anos no primeiro e 24 anos no segundo, evidenciando uma falta de isonomia nos grupos. Os resultados da pesquisa são apresentados na tabela 10 (o nível de referência é “15-25 anos”).

| VD | | | | |
|--|-----------------|---------------|----------|--------------------|
| <i>Predictors</i> | <i>Log-Odds</i> | <i>CI</i> | <i>p</i> | <i>Response</i> |
| (Intercept) | -0.71 | -0.93 – -0.48 | <0.001 | Dupla Negação |
| faixa.etaria26-50 anos | -0.65 | -0.97 – -0.34 | <0.001 | Dupla Negação |
| (Intercept) | -4.36 | -5.50 – -3.22 | <0.001 | Negação Pós-verbal |
| faixa.etaria26-50 anos | 0.72 | -0.58 – 2.02 | 0.278 | Negação Pós-verbal |
| Observations | 843 | | | |
| R ² / R ² adjusted | 0.017 / 0.015 | | | |

Tabela 10. Influência da variável “faixa etária” em Nunes (2014). Fonte: Santos (2023).

Nessa pesquisa, a faixa etária foi um fator significativo no decréscimo do uso da Dupla Negação. A alteração da faixa etária mais jovem para a mais adulta leva os *log-odds* de -0,71 (32,9%) para -1,36 (20,4%). Ou seja, a probabilidade de uso da Dupla Negação em um falante entre 26 e 50 anos é significativamente menor do que em um falante entre 15 e 25 anos. Mesmo assim, o R² ajustado ainda é muito baixo (0,015), revelando que essa variável explica muito pouco da variabilidade dos dados, embora seja um número maior que o das outras pesquisas.

Por fim, Nascimento (2014) analisa a faixa etária a partir dos níveis “15-25 anos”, “26-49 anos” e “>50 anos”. Mais uma vez, falta isonomia na divisão das células, dificultando a generalização dos dados. Não só isso, mas nessa pesquisa, ao contrário de Nunes (2014), há um nível a mais para a faixa etária, o que prejudica a comparação. Os resultados estão dispostos na tabela 11.

| VD | | | | |
|--|-----------------|---------------|----------|--------------------|
| <i>Predictors</i> | <i>Log-Odds</i> | <i>CI</i> | <i>p</i> | <i>Response</i> |
| (Intercept) | -1.24 | -1.41 – -1.06 | <0.001 | Dupla Negação |
| faixa.etaria26-49 anos | -0.10 | -0.35 – 0.15 | 0.435 | Dupla Negação |
| faixa etaria [>49 anos] | -0.08 | -0.32 – 0.16 | 0.520 | Dupla Negação |
| (Intercept) | -3.63 | -4.14 – -3.12 | <0.001 | Negação Pós-verbal |
| faixa.etaria26-49 anos | -0.75 | -1.65 – 0.16 | 0.106 | Negação Pós-verbal |
| faixa etaria [>49 anos] | -0.33 | -1.10 – 0.44 | 0.404 | Negação Pós-verbal |
| Observations | 2263 | | | |
| R ² / R ² adjusted | 0.001 / 0.001 | | | |

Tabela 11. Influência da variável “faixa etária” em Nascimento (2014). Fonte: Santos (2023).

Nesse estudo, a VI “faixa etária” só se mostrou significativa no Intercepto, com R^2 próximo de zero e intervalos de confiança cruzando o zero, indicando uma associação praticamente nula. Assim, percebe-se que a comunidade do Espírito Santo está em variação estável entre as formas de negação, ou seja, falantes mais jovens e mais idosos as usam da mesma maneira.

Os estudos, portanto, não apresentam convergência nos seus resultados. Nos estudos do Nordeste, Cavalcante (2007) reportou significância no desfavorecimento da Negação Pós-verbal com o aumento da idade e insignificância na Dupla Negação, enquanto Lopes da Silva (2020) reportou significativo desfavorecimento no uso da Dupla Negação em Fortaleza (CE). Para explicar essa diferença, afora a já conhecida distinção entre as comunidades da Bahia e do Ceará, percebe-se uma diferença no modo como seccionaram a própria VI. Se não há isonomia no modo como ela é estruturada, os resultados podem ser distintos – como foi o caso.

Da mesma maneira, no Sudeste, analisado em Rocha (2013), Nunes (2014) e Nascimento (2014), os resultados não foram convergentes. Em Rocha (2013), que analisou a comunidade de São Paulo, a faixa etária 3 favorece significativamente a Dupla Negação. Já em Nunes (2014), que analisou a comunidade do Rio de Janeiro, a Dupla Negação foi significativamente desfavorecida entre os mais velhos, e a Negação Pós-verbal não foi influenciada pela faixa etária maior. Por fim, Nascimento (2014) mostrou insignificância da faixa etária no uso dos dois tipos de negação.

Nesse sentido, observa-se como o fenômeno da negação não é convergente nem no Nordeste, nem no Sudeste. Ao comparar os resultados das duas regiões, percebe-se que houve um desfavorecimento geral dos dois tipos de negação quando comparados à Negação Pré-verbal no Nordeste, semelhante ao que ocorre no Rio de Janeiro, e contrário ao que ocorre em São Paulo. Contudo, vale ressaltar que uma comparação mais profunda é impossibilitada pela falta de isonomia na definição de tal VI, o que será discutido com mais detalhes na seção seguinte.

6. CONSIDERAÇÕES FINAIS

Esta pesquisa objetivou investigar a relevância estatística dos condicionamentos sociais comumente utilizados em pesquisas na Sociolinguística Variacionista: sexo/gênero, escolaridade e faixa etária. Para isso, utilizou-se como variável dependente um fenômeno linguístico variável de âmbito morfossintático: a negação verbal. Como recorte de pesquisa, foram selecionadas as regiões Nordeste e Sudeste pelo fato de as duas regiões apresentarem fluxos migratórios entre si ao longo dos anos (Villa, 2017). Assim, definiu-se o tipo de pesquisa: uma meta-análise, o meio mais “poderoso” dentro da ciência para se comparar e sintetizar o resultado de trabalhos distintos sobre um mesmo tema.

Primeiramente, notou-se uma prevalência em todas as regiões da Negação Pré-verbal, seguida da Dupla Negação e da Negação Pós-verbal. Embora se tenha notado um favorecimento da Negação

Pós-verbal no Nordeste, especificamente em Cavalcante (2007), no Ceará e no Maranhão não houve favorecimento dessa variante com a mesma intensidade, enquanto em alguns dos trabalhos no Sudeste, sim. Nesse sentido, com bases nesses dados, não parece ser o contato linguístico o diferencial entre as duas regiões para o uso da Negação Pós-verbal, ou então os fluxos migratórios ao longo do século XX (Villa, 2017) acabaram levando a variante para as variedades do Sudeste.

Acerca das variáveis investigadas, notou-se um comportamento divergente. Enquanto o sexo/gênero do falante não parece ser um fator influente na variação entre as formas de negação, a escolaridade parece ser um condicionamento parcial, sendo mais influente nos anos iniciais e finais do ensino fundamental, tornando-se insignificante se o falante ingressar no ensino superior. Por fim, a faixa etária apresentou resultados divergentes, com alguns trabalhos identificando um consistente desfavorecimento da Dupla Negação no Sudeste (Nunes, 2014) e uma pequena valorização no Ceará (Lopes da Silva, 2020) – este com intervalo de confiança próximo do zero em seu limite inferior e um R^2 próximo a zero, indicando uma baixa relevância da variável.

Embora a quantidade de trabalhos seja baixa, esta meta-análise poderia ser expandida em outras pesquisas utilizando todos os estudos sociolinguísticos sobre negação verbal no português. Além disso, também seria interessante novas meta-análises investigarem a relevância dos tradicionais condicionamentos sociais utilizados em outros fenômenos morfossintáticos, a fim de se entender se já não é hora de deixá-los de lado em busca de outras pressões sociais que sejam mais relevantes nesse nível de análise linguística. A hipótese inicial do trabalho sobre a insignificância das VIs, portanto, não pode ser confirmada em sua totalidade, porque há divergências estruturais entre as pesquisas, impossibilitando uma generalização dos resultados.

Por fim, recomenda-se a realização de novas meta-análises, pelo menos em trabalhos nos grandes centros urbanos, para que a crescente literatura em Sociolinguística Variacionista seja sistematizada e as generalizações que a área propõe sejam alcançadas. Sem esse mecanismo de análise, os pesquisadores ficarão reféns de revisões de literatura narrativas, que têm menos poder explanatório e podem ser mais enviesadas, já que muitos trabalhos sobre o tema acabam ficando de fora por não terem sido sistematizados antes.

INFORMAÇÕES COMPLEMENTARES

CONFLITO DE INTERESSE

Os autores não têm conflitos de interesse a declarar.

DECLARAÇÃO DE DISPONIBILIDADE DE DADOS

Os dados, códigos e materiais que suportam os resultados deste estudo estão disponíveis abertamente no GitHub através do link: <https://github.com/phsousaa/Data-and-scripts>.

DECLARAÇÃO DE USO DE IA

Os autores declaram que nenhuma ferramenta de IA foi utilizada na criação deste manuscrito nem em qualquer aspecto dos trabalhos realizados cujo resultado está reportado no manuscrito.

AValiação E RESPOSTA DOS AUTORES

Avaliação: <https://doi.org/10.25189/2675-4916.2025.V6.N4.ID858.R>

Resposta dos Autores: <https://doi.org/10.25189/2675-4916.2025.V6.N4.ID858.A>

REFERÊNCIAS

ALMEIDA, Wellington Couto de. *Processamento e Percepção da Concordância Verbal Variável de P6 entre Universitários da Cidade do Rio de Janeiro*. 2022. Dissertação (Mestrado em Linguística) – Departamento de Linguística e Filologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2022.

BECHARA, Evanildo. *Moderna Gramática Portuguesa*. Rio de Janeiro: Nova Fronteira, 2019.

BRAGA, Luciana; SILVA, Josenildo Ferreira Teófilo da. Análise Funcionalista das Estratégias de Negação do Português Oral Culto de Fortaleza: um estudo de caso. *Entrepalavras*, v. 1, n. 1, p. 69-84, 2011.

BORTONI-RICARDO, Stella Maris. *Educação em língua materna: a sociolinguística na sala de aula*. São Paulo: Parábola, 2004.

CAMPBELL-KIBLER, Kathryn. The nature of sociolinguistic perception. *Language Variation and Change*, v. 21, n. 1, p. 135-156, mar. 2009. DOI: <https://doi.org/10.1017/S0954394509000052>.

CASTRO, Savanna Souza de; VIEIRA, Julinara Silva; SOUSA, Valéria Viana; SILVA, Jorge Augusto Alves. Uma Análise Sociofuncionalista da Dupla Negação no Sertão da Ressaca. In: IX Seminário de Pesquisa e Estudos Linguísticos, 9., 2017, Vitória da Conquista. *Anais [...]*. Vitória da Conquista: UESB, 2017. p. 143-148.

CASTILHO, Ataliba de. *Nova gramática do português brasileiro*. São Paulo: Contexto, 2019.

CAVALCANTE, Rerisson. *A Negação Pós-verbal no Português Brasileiro: análise descritiva e teórica de dialetos rurais de afro-descendentes*. Dissertação (Mestrado em Letras) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2007.

CAVALCANTI, Jauranice Rodrigues. As faces de uma polêmica: o episódio do livro didático Por uma vida melhor. *D.E.L.T.A.*, v. 29, n. especial, p. 485-501, 2013.

COELHO, Izete Lehmkuhl, Görski, Edair Maria; May, Guilherme Henrique; Souza, Christiane Maria Nunes. *Sociolinguística*. Florianópolis: UFSC, 2010.

CUNHA, Angélica Furtado da. Funcionalismo. In: MARTELOTTA, Mário Eduardo (org.). *Manual de Linguística*. 2. ed. 10. reimpr. São Paulo: Contexto, 2023. p. 157-176.

ECKERT, Penelope; LABOV, William. Phonetics, Phonology and Social Meaning. *Journal of Sociolinguistics*, v. 21, n. 4, p. 1-30, set. 2017. DOI: <https://doi.org/10.1111/josl.12244>.

FREITAG, Raquel M. K.; PINHEIRO, Bruno Felipe Marques. Modelo de Árvore de Inferência Condicional para Explicar Usos Linguísticos Variáveis. In: CARVALHO, Cristina dos Santos; LOPES, Norma da Silva; RODRIGUES, Angélica (org.). *Sociolinguística e Funcionalismo: vertentes e interfaces*. Salvador: EDUNEB, 2020. p. 317-342.

FREITAG, Raquel M. K. *Como fazer meta-análise com dados sociolinguísticos?*. Disponível em: <https://rkofreitag.github.io/meta.html>. Atualizado em: 2021-04-11. Acesso: 31 out. 2025.

GIUSTINI, Dean; BOULOS, Maged N. Kamel. Google Scholar is not enough to be used alone for systematic reviews. *Online Journal of Public Health Informatics*, v. 5, n. 2, p. 1-10, 2013. DOI: <https://doi.org/10.5210%2Fojphi.v5i2.4623>.

GRIES, Stefan. *Estatística com R para a Linguística: uma introdução prática*. Belo Horizonte: FALE/UFMG, 2019.

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

LABOV, William. Where does the Linguistic Variable Stop? A response to Beatriz Lavandera. *Working Papers in Sociolinguistics*, n. 44, p. 1-23, 1978. Disponível em: <https://eric.ed.gov/?id=ED157378>. Acesso em: 28 fev. 2025.

LABOV, William. *Padrões Sociolinguísticos*. São Paulo: Parábola, 2008.

LAVANDERA, Beatriz. Where Does the Sociolinguistic Variable Stop? *Language in Society*, v. 7, n. 2, p. 171-182, 1978. DOI: <https://doi.org/10.1017/S0047404500005510>.

LEE, Dong Kyu. Alternatives to P value: confidence interval and effect size. *Korean Journal of Anesthesiology*, v. 69, n. 6, p. 555-562, 2016. DOI: <https://doi.org/10.4097/kjae.2016.69.6.555>.

LOPES DA SILVA, Késsio Jhone. A Variação entre a Dupla Negação e a Negação Pré-verbal no Português de Fortaleza-CE. *Letra Magna*, v. 16, n. 25, p. 1279-1299, 2020.

LOVATTO, Paulo Alberto; Lehnen, Cheila Roberta; Andretta, Ines; CARVALHO, Anderson de; HAUSCHILD, Luciano. Meta-análise em Pesquisas Científicas: enfoque em metodologias. *Revista Brasileira de Zootecnia*, v. 36, p. 285-294, jul. 2007. DOI: <https://doi.org/10.1590/S1516-35982007001000026>.

LÜDECKE, Daniel. sjPlot: Data Visualization for Statistics in Social Science. *CRAN*, 2023. R package version 2.8.14. Disponível em: <https://CRAN.R-project.org/package=sjPlot>. Acesso em: 9 abr. 2023.

MIOTO, Carlos; SILVA, Maria Cristina Figueiredo; LOPES, Ruth. *Novo Manual de Sintaxe*. São Paulo: Contexto, 2018.

MOLLICA, Maria Cecília; BRAGA, Maria Luiza. *Introdução à Sociolinguística: o tratamento da variação*. 4. ed. São Paulo: Contexto, 2020.

NARO, Anthony Julius. Modelos quantitativos e tratamento estatístico. In: MOLLICA, Maria Cecília; BRAGA, Maria Luiza (Org.). *Introdução à Sociolinguística: o tratamento da variação*. 4. ed. São Paulo: Contexto, 2020. p. 15-25.

NASCIMENTO, Cristina Aparecida Reimann do. *A Negação no Português Falado em Vitória/ES*. 2014. Dissertação (Mestrado em Linguística) – Universidade Federal do Espírito Santo, Vitória, 2014.

NOVAIS, Viviane; SIQUEIRA, Manoel. A Variável Sexo/gênero no Português Falado no Sertão Alagoano. *Leitura*, Maceió, n. 66, p. 35-50, 2020.

NUNES, Elizene Sebastiana de Oliveira. A Negação no Português Falado do Rio de Janeiro: um estudo baseado em *corpus*. *Revista do SELL*, v. 4, n. 1, p. 1-19, 2014.

OUSHIRO, Livia. Introdução à Estatística para Linguistas. *Zenodo*, 2017. DOI: <https://doi.org/10.5281/zenodo.4755739>.

PATIL, Indrajeet. Visualizations with statistical details: The 'ggstatsplot' approach. *Journal of Open Source Software*, v. 6, n. 61, p. 1-5, 2021. DOI: 10.21105/joss.03167.

PERINI, Mario A. *Gramática Descritiva do Português Brasileiro*. Petrópolis: Vozes, 2016.

R CORE TEAM. *R: A language and environment for statistical computing*. Áustria: R Foundation for Statistical Computing, 2021. Disponível em: <https://www.R-project.org/>. Acesso em: 31 out. 2025.

REIMANN, Cristiana Aparecida; YACOVENCO, Lilian Coutinho. A Dupla Negação no Português Falado em Vitória/ES: traço da identidade linguística capixaba?. In: CONGRESSO NACIONAL DE ESTUDOS LINGÜÍSTICOS, 1., 2011, Vitória. *Anais [...]*. Vitória: UFES, 2011. p. 1-4.

REIMANN, Cristiana Aparecida; YACOVENCO, Lilian Coutinho. A Negação no Português Falado em Vitória/ES: atuação de fatores discursivos e sintáticos. In: CONGRESSO NACIONAL DE ESTUDOS LINGÜÍSTICOS, 2., 2011, Vitória. *Anais [...]*. Vitória: UFES, 2011. p. 33-37.

ROCHA, Rafael Stoppa. Negação Verbal no Português Paulistano: envelope de variação. *Estudos Linguísticos*, v. 41, n. 2, p. 833-843, 2012.

ROCHA, Rafael Stoppa. *A Negação Dupla no Português Paulistano*. Dissertação (Mestrado em Linguística) – Departamento de Linguística, Universidade de São Paulo, São Paulo, 2013.

SANTOS, Pedro Henrique Sousa dos. *Meta-análise dos estudos de negação verbal nas regiões Nordeste e Sudeste: investigação da relevância dos condicionamentos sociais*. Trabalho de Conclusão de Curso (Licenciatura em Letras – Língua Portuguesa) – Faculdade de Letras, Universidade Federal de Alagoas, Arapiraca, 2023.

SCHERRE, Maria Marta Pereira. Sobre a Influência de Variáveis Sociais na Concordância Nominal. In: OLIVEIRA E SILVA, Giselle Machline; SCHERRE, Maria Marta Pereira (Orgs). *Padrões Sociolinguísticos: análise de fenômenos variáveis no português falado na cidade do Rio de Janeiro*. Rio de Janeiro: Tempo Brasileiro, 1996. p. 85-117.

SCHWENTER, Scott A. The Pragmatics of Negation in Brazilian Portuguese. *Lingua*, v. 115, n. 10, p. 1427-1456, 2005. DOI: <https://doi.org/10.1016/j.lingua.2004.06.006>.

SERRA, Flávia Pereira. *“Eu Não Digo ‘Não’ Duas Vezes ‘Não’”: usos e percepções avaliativas sobre a Dupla Negação no português falado no Maranhão*. 2018. Dissertação (Mestrado em Letras) – Centro de Ciências Humanas, Universidade Federal do Maranhão, São Luís, 2018.

SILVA, Vera Lúcia Paredes da. Relevância das variáveis linguísticas. In: MOLLICA, Maria Cecília; BRAGA, Maria Luíza (Orgs). *Introdução à Sociolinguística: o tratamento da variação*. 4. ed. São Paulo: Contexto, 2020. p. 67-71.

SIQUEIRA, Manoel. Análise Contrastiva da Estrutura do Sintagma Nominal Possessivizado no Português Brasileiro. *Matraga*, v. 28, n. 52, p. 25-43, 2021. DOI: 10.12957/matraga.2021.53146.

VILLA, Marco Antonio. *Quando eu vim-me embora: história da migração nordestina para São Paulo*. Rio de Janeiro: LeYa, 2017.

VITÓRIO, Elyne Giselle de Santana Lima Aguiar. A Língua Usada no Sertão Alagoano: constituição da amostra. In: OLIVEIRA, Almir Almeida; PAULA, Aldir Santos. (Orgs.). *Interfaces Sociolinguísticas: análises variacionistas em Alagoas e Pernambuco*. Arapiraca: EdUNEAL, 2020.

VITÓRIO, Elyne Giselle de Santana Lima Aguiar. A difusão de *a gente* sujeito na variedade alagoana: um estudo de meta-análise. *Revista Diadorim*, v. 26, n. 2, p. 1-21, 2024. Disponível em: <https://revistas.ufrj.br/index.php/diadorim/article/view/64304>. Acesso em: 7 mar. 2025.

VENABLES, William Norman; RIPLEY, Brian. *Modern Applied Statistics with S*. 4. ed. New York: Springer, 2002.

WICKHAM, Hadley; AVERICK, Mara; BRYAN, Jennifer; CHANG, Winston; D’AGOSTINO MCGOWAN, Lucy; FRANÇOIS, Romain; GROLEMUND, Garrett; HAYES, Alex; HENRY, Lionel; HESTER, Jim; KUHN, Max; LIN PEDERSEN, Thomas; MILLER,

Evan; MILTON BACHE, Stephan; MÜLLER, Kirill; OOMS, Jeroen; ROBINSON, David; PAIGE SEIDEL, Dana; SPINU, Vitalie; TAKAHASHI, Kohske; VAUGHAN, Davis; WILKE, Claus; WOO, Kara; YUTANI, Hiroaki. Welcome to the Tidyverse. *Journal of Open Source Software*, v. 4, n. 43, p. 1-6, 2019. DOI: 10.21105/joss.01686.

YACOVENCO, Lilian Coutinho; NASCIMENTO, Cristina Aparecida Reimann do. A Negação no Português Falado em Vitória/ES. *Revista (Con)Textos Linguísticos*, v. 10, n. 17, p. 122-138, 2016.