



**UNIVERSIDADE FEDERAL DE ALAGOAS – UFAL**  
**CAMPUS ARAPIRACA**  
**GRADUAÇÃO EM LETRAS – LÍNGUA PORTUGUESA**

**PEDRO HENRIQUE SOUSA DOS SANTOS**

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

**ARAPIRACA**

**2023**

Pedro Henrique Sousa dos Santos

**Meta-análise dos estudos de negação verbal nas regiões Nordeste e Sudeste: investigação da relevância dos condicionamentos sociais**

Monografia apresentada como requisito parcial para a obtenção do grau de Licenciado em Letras, pela Universidade Federal de Alagoas – UFAL, *campus* de Arapiraca.

Orientadora: Profa. Dra. Elyne Giselle de Santana Lima Aguiar Vitória

Arapiraca

2023



Universidade Federal de Alagoas – UFAL  
Campus Arapiraca  
Biblioteca Setorial *Campus Arapiraca* - BSCA

S237p Santos, Pedro Henrique Sousa dos  
Meta-análise dos estudos de negação verbal nas regiões Nordeste e Sudeste  
[recurso eletrônico]: investigação da relevância dos condicionamentos sociais / Pedro  
Henrique Sousa dos Santos. – Arapiraca, 2023.  
71 f.: il.

Orientadora: Profa. Dra. Elyne Giselle de Santana Lima Aguiar Vitória.  
Trabalho de Conclusão de Curso (Licenciatura em Letras – Língua Portuguesa) -  
Universidade Federal de Alagoas, *Campus Arapiraca*, Arapiraca, 2023.  
Disponível em: Universidade Digital (UD) – UFAL (*Campus Arapiraca*).  
Referências: f. 45-51.  
Apêndices: f. 52-71

1. Letras. 2. Sociolinguística. 3. Variação linguística. I. Vitória, Elyne Giselle de  
Santana Lima Aguiar. II. Título.

CDU 81


Pedro Henrique Sousa dos Santos

Meta-análise dos estudos de negação verbal nas regiões nordeste e sudeste: investigação da relevância dos condicionamentos sociais

Trabalho de conclusão de Curso apresentado ao Curso de Letras – Língua Portuguesa da Universidade Federal de Alagoas – UFAL, *Campus* Arapiraca, como requisito parcial para obtenção do grau de Licenciado em Letras – Língua Portuguesa.


Data da aprovação: 15 / 09 / 2023

#### Banca Examinadora

Documento assinado digitalmente  
 ELYNE GISELLE DE SANTANA LIMA AGUIAR VITO  
Data: 16/09/2023 05:54:58-0300  
Verifique em <https://validar.iti.gov.br>


---

Profa. Dra. Elyne Giselle de Santana Lima Aguiar Vitória  
Universidade Federal de Alagoas – UFAL  
*Campus* Arapiraca  
(Orientador)

Documento assinado digitalmente  
 JOSE MANOEL SIQUEIRA DA SILVA  
Data: 16/09/2023 20:38:38-0300  
Verifique em <https://validar.iti.gov.br>

---

Prof. Me. José Manoel Siqueira da Silva  
Universidade Federal de Sergipe – UFS  
PPGL  
(Examinador)

Documento assinado digitalmente  
 ELIANE VITORINO DE MOURA OLIVEIRA  
Data: 17/09/2023 17:46:15-0300  
Verifique em <https://validar.iti.gov.br>

---

Profa. Dra. Eliane Vitorino de Moura Oliveira  
Universidade Federal de Alagoas – UFAL  
*Campus* Arapiraca  
(Examinador)

## 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 uma sistematização das pesquisas sobre cada tema a fim de se encontrar possíveis lacunas e tecer generalizações, adentra-se, a partir disso, a revisão de literatura. Entre elas, a meta-análise é considerada a com maior poder explicativo (Freitag, 2021). Nesse contexto, tomando como base um fenômeno morfossintático, a negação verbal, que varia entre as formas “negação pré-verbal”, “dupla negação” e “negação pós-verbal”, este trabalho objetivou investigar a relevância de três condicionamentos sociais em uma meta-análise, a saber: sexo/gênero, escolaridade e faixa etária. Essas restrições foram escolhidas por serem comuns nos estudos variacionistas (Mollica; Braga, 2020) e porque, nos recentes estudos de fenômenos morfossintáticos, muitas vezes não apresentam significância estatística, como foi o caso da variável “sexo/gênero” em Novais e Siqueira (2020). Para isso, foi preciso delimitar a análise a duas regiões, Nordeste e Sudeste, por conta da inter-relação histórica entre elas (Villa, 2017) e por conta do escopo do trabalho — uma monografia, cujo espaço de tempo e características textuais não permitem uma ampla discussão. Como resultados desta meta-análise, foram percebidas baixa ou nula influência de “sexo/gênero”, significativa influência de “escolaridade” quando o nível “ensino fundamental” é alterado para “ensino médio” (o ensino superior não se mostrou um nível significativo) e baixa ou nula influência de “faixa etária”, indicando uma estabilidade das formas. Depois dos resultados, ainda foram discutidos alguns problemas relativos à feitura da meta-análise, desde a seleção dos dados até a organização das variáveis em cada pesquisa. Nesse sentido, o trabalho contribuiu para a ampliação de trabalhos 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 mais bem sistematizados.

**Palavras-chave:** sociolinguística; variação; negação verbal; meta-análise; condicionamento social.

## 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 works at the most diverse grammatical levels, Sociolinguistics begins to need a systematization of research on each topic in order to find possible gaps and make generalizations, with the use of literature review. Among them, meta-analysis is considered the one with the greatest explanatory power (Freitag, 2021). In this context, based on a morphosyntactic phenomenon, verbal negation, which varies between the forms “pre-verbal negation”, “double negation” and “post-verbal negation”, this work aimed to investigate the relevance of three social conditionings in a meta-analysis, namely: sex/gender, education, and age group. These restrictions were chosen because they are common in variationist studies (Mollica; Braga, 2020) and because, in recent studies of morphosyntactic phenomena, they often do not present statistical significance, as was the case with the “sex/gender” variable in Novais and Siqueira (2020). To achieve this, it was necessary to segment the analysis to two regions, Northeast and Southeast, due to the historical interrelationship between them (Villa, 2017) and due to the scope of the work — a monograph, whose time frame and textual characteristics do not allow a broad discussion. As a result of this meta-analysis, low or no influence of “sex/gender” was perceived, as well as a significant influence of “education” when the level of “elementary education” is changed to “secondary education” (higher education did not prove to be a significant level) and low or no influence of “age group”, indicating a stability of the forms. After the results, some problems related to carrying out the meta-analysis were discussed, from data selection to the organization of variables in each research. In this sense, the work contributed to the expansion of studies that deal with meta-analysis in Variationist Sociolinguistics, pointing to parameters 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.

## LISTA DE TABELAS

|   |    |
|---|----|
| Tabela 1 – Distribuição da negação em pesquisas selecionadas pela revisão sistemática ..... | 27 |
| Tabela 2 – Influência da variável “escolaridade” em Cavalcante (2007) .....                 | 31 |
| Tabela 3 – Influência da variável “escolaridade” em Lopes da Silva (2020) .....             | 32 |
| Tabela 4 – Influência da variável “escolaridade” em Rocha (2013) .....                      | 33 |
| Tabela 5 – Influência da variável “escolaridade” em Nunes (2014) .....                      | 34 |
| Tabela 6 – Influência da variável “escolaridade” em Nascimento (2014) .....                 | 34 |
| Tabela 7 – Influência da variável “faixa etária” em Cavalcante (2007) .....                 | 36 |
| Tabela 8 – Influência da variável “faixa etária” em Lopes da Silva (2020) .....             | 37 |
| Tabela 9 – Influência da variável “faixa etária” em Rocha (2013) .....                      | 38 |
| Tabela 10 – Influência da variável “faixa etária” em Nunes (2014) .....                     | 39 |
| Tabela 11 – Influência da variável “faixa etária” em Nascimento (2014) .....                | 39 |

## SUMÁRIO

|              |  |           |
|--------------|--|-----------|
| <b>1</b>     | <b>INTRODUÇÃO</b>  | <b>7</b>  |
| <b>2</b>     | <b>A NEGAÇÃO NO PORTUGUÊS BRASILEIRO</b>                                   | <b>9</b>  |
| <b>2.1</b>   | <b>A Negação</b>   | <b>9</b>  |
| <b>3</b>     | <b>FUNDAMENTAÇÃO TEÓRICA</b>   | <b>13</b> |
| <b>3.1</b>   | <b>Sociolinguística Variacionista</b>                                      | <b>13</b> |
| <b>3.2</b>   | <b>Variáveis Sociais: sexo/gênero, faixa etária e escolaridade</b>         | <b>14</b> |
| <b>3.2.1</b> | <b>Sexo/gênero</b>   | <b>14</b> |
| <b>3.2.2</b> | <b>Faixa etária</b>  | <b>16</b> |
| <b>3.2.3</b> | <b>Escolaridade</b>  | <b>17</b> |
| <b>4</b>     | <b>METODOLOGIA</b>   | <b>19</b> |
| <b>4.1</b>   | <b>Ferramentas Estatísticas Utilizadas: R e RStudio</b>                    | <b>19</b> |
| <b>4.2</b>   | <b>Meta-análise</b>  | <b>20</b> |
| <b>4.3</b>   | <b>Medidas Estatísticas</b>  | <b>21</b> |
| <b>5</b>     | <b>RESULTADOS</b>  | <b>24</b> |
| <b>5.1</b>   | <b>Revisão Sistemática de Literatura</b>                                   | <b>24</b> |
| <b>5.2</b>   | <b>Análise Geral</b>   | <b>26</b> |
| <b>5.3</b>   | <b>Análise da Variável Sexo/gênero</b>                                     | <b>28</b> |
| <b>5.4</b>   | <b>Análise da Variável Escolaridade</b>                                    | <b>31</b> |
| <b>5.5</b>   | <b>Análise da Variável Faixa Etária</b>                                    | <b>35</b> |
| <b>6</b>     | <b>PROBLEMAS EM META-ANÁLISES DENTRO DA SOCIOLINGUÍSTICA VARIACIONISTA</b> | <b>41</b> |
| <b>7</b>     | <b>CONSIDERAÇÕES FINAIS</b>  | <b>43</b> |
|              | <b>REFERÊNCIAS</b>   | <b>45</b> |
|              | <b>APÊNDICE A – PACOTES UTILIZADOS E ANÁLISE GERAL DOS DADOS</b>           | <b>52</b> |
|              | <b>APÊNDICE B – ANÁLISE DA VARIÁVEL SEXO/GÊNERO</b>                        | <b>53</b> |
|              | <b>APÊNDICE C – ANÁLISE DA VARIÁVEL ESCOLARIDADE</b>                       | <b>62</b> |
|              | <b>APÊNDICE D – ANÁLISE DA VARIÁVEL FAIXA ETÁRIA</b>                       | <b>67</b> |



## 1 INTRODUÇÃO

Este estudo tem como objetivo identificar a relevância dos condicionamentos sociais no uso da negação verbal no Nordeste e no Sudeste por meio de uma 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, é utilizado o fenômeno da negação verbal, que é constituído pelas variantes “Negação Pré-verbal”, “Negação Pós-verbal” e “Dupla Negação”, como em (1a), (1b) e (1c)<sup>1</sup>. Como recorte de pesquisa, são selecionadas as regiões Nordeste e Sudeste, tanto por o escopo do trabalho ser menor — uma monografia — quanto pelo fato de as duas regiões apresentarem inter-relação histórica devido às migrações humanas durante o século XX (Villa, 2017). Outro motivo é que a região Nordeste é a localização da Universidade a que este trabalho é vinculado, o que torna necessária uma análise de fatores locais em comparação a outras regiões. Assim, pretende-se investigar se a negação é influenciada pelas variáveis sociais usualmente escolhidas para estratificação de amostras: sexo/gênero, faixa etária e escolaridade.

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

Tais variáveis são padrões na Sociolinguística Variacionista do Brasil para a constituição das células sociais, mas, principalmente em fenômenos morfossintáticos<sup>2</sup>, sua correlação com os usos é, em algumas pesquisas, não significativa, de um ponto de vista estatístico (cf. Novais; Siqueira, 2020). Nesse sentido, é importante investigar se a negação é ou não influenciada por elas, o que pode impactar futuras pesquisas sobre o fenômeno, que não seccionariam a população em um desses fatores caso não seja significativo no condicionamento da variação. Em outras palavras, caso a escolaridade, por exemplo, não seja condicionante de

<sup>1</sup>Exemplos criados com base em introspecção.

<sup>2</sup>No caso de fenômenos como a concordância, há efeitos da prescrição sobre o uso, que se reflete na significância estatística de variáveis como a escolaridade.

um dos usos da negação, então futuras pesquisas não levariam em conta essa variável, visto que pessoas de baixa e alta escolaridades usariam a negação da mesma maneira.

Para isso, foram elencados os seguintes objetivos específicos: i) delimitar os descritores para a busca dos trabalhos; ii) delimitar a base de dados em que será feita a busca pelos trabalhos sobre negação verbal; iii) realizar uma revisão sistemática da literatura sobre o fenômeno abordado; iv) realizar a meta-análise pela aplicação de testes estatísticos aos dados dos textos selecionados; v) comparar, quando possível, os dados dos trabalhos; e vi) descrever a influência dos condicionamentos sociais no uso da negação verbal.

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 dois, sintetiza-se o fenômeno da negação no português brasileiro mediante distintas perspectivas de análise. Em seguida, na seção três, destaca-se a fundamentação teórica para o trabalho, delimitando os termos referentes à área de estudo e às variáveis independentes investigadas. Já a quarta seção trata da metodologia de pesquisa, descrevendo as ferramentas estatísticas utilizadas, os procedimentos de realização do trabalho e a descrição de como a meta-análise foi realizada. Após isso, são descritos os resultados tanto de um modo geral, para analisar a negação verbal nas duas regiões, quanto de modo específico, com uma subseção para cada condicionamento social analisado. Depois, há uma seção sintetizando os problemas em se fazer uma meta-análise em Sociolinguística e alguns princípios para futuras pesquisas seguirem e facilitarem, desse modo, a sistematização. Por fim, apresentam-se as considerações finais, reportando as contribuições e as limitações do presente estudo para a Sociolinguística Variacionista.

## 2 A NEGAÇÃO NO PORTUGUÊS BRASILEIRO

Nesta seção, é explorado o modo como é descrita a negação em diversos trabalhos. Partindo tanto de gramáticas normativas quanto descritivas e chegando a estudos linguísticos formais e funcionais, dá-se um enfoque ao modo como esse fenômeno é complexo e necessita de mais trabalhos para caracterizá-lo em sua totalidade.

### 2.1 A Negação

A negação é um fenômeno recorrente nas línguas humanas, ocorrendo no português brasileiro principalmente por meio da partícula “não” pré-verbal, como em (2) — 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*.

Além disso, conforme o gramático, a negação atua sobre o significado ôntico (particular ao ser) das orações, visto que pode alterar o valor de existência destas de modo a torná-lo falso ou duvidoso. Esse mesmo caráter de alteração global das orações faz com que Rocha Lima (2011, p. 229) considere a negação como uma palavra *denotativa* (tendo significado próprio), e não como um advérbio, já que não há, de fato, expressão de circunstância.

Tais regras acrescem-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 em (4), e, quando há processos de subordinação, esse *não* final nega toda a sentença, e não apenas a oração subordinada, como em (5), exceto em casos de coordenação, se nos quais o segundo *não* vier posposto ao verbo da primeira oração, como em (6)<sup>3</sup>.

(4) Eu não vou lá *não*.

---

<sup>3</sup>Exemplos retirados de Perini (2016, p. 171-172).

- (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.

Com essas descrições, percebe-se o escopo da negação verbal, que é objeto do presente estudo. Entretanto, por motivos de esclarecimento, ressalta-se que Perini (2016, p. 173) também comenta sobre a negação nominal e adverbial, em que o escopo da negação atua sobre tais partículas, como em (7) e (8) — exemplos de Perini (2016). Por não ser verbal, não será descrito aqui, embora se deva saber que há alta produtividade de tal fenômeno no português brasileiro.

- (7) Os candidatos *não aprovados* podem tentar novamente.
- (8) Você pode tocar bateria, mas *nunca aqui*.

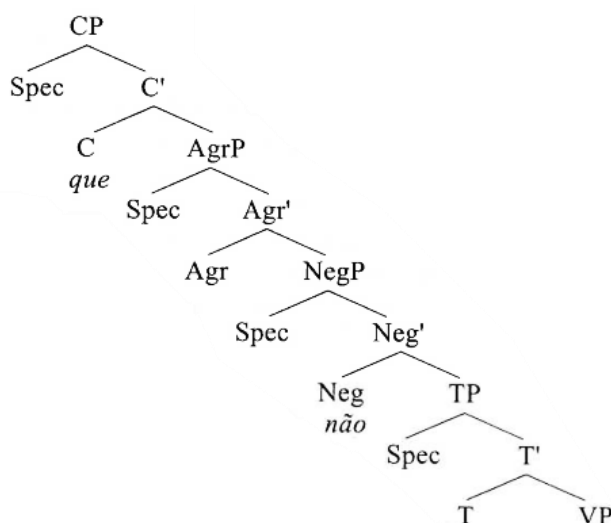
Acerca da negação, Castilho (2019, p. 577) descreve quatro tipos: a negação anterior ao verbo, funcionando como um quase clítico, a negação pós-verbal, a negação no final da sentença (funcionando como uma pergunta) e a cliticização do *não* em um quase prefixo, como em (9), (10), (11) e (12), exemplos retirados da obra do autor.

- (9) A programação havia sido planejada, mas *não* deu certo.
- (10) Tá gostando *não*?
- (11) A menina toma conta – precocemente... *não*? – das atividades dos irmãos...
- (12) Atividades *não* lucrativas.

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. Mioto, 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), conforme Figura 1.

Figura 1 – Representação da posição estrutural de NegP



Fonte: Mioto, Silva e Lopes (2018, p. 64).

Assim como os outros sintagmas, NegP domina<sup>4</sup> o sintagma verbal, gravitando em torno dele. Os autores também comentam sobre a natureza funcional dessa categoria, que seleciona categorialmente (c-seleção) outros elementos como complementos ou especificadores, impossibilitados de fazer seleção semântica (s-seleção), 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 *Não* pré-verbal funciona como núcleo e sua redução fonética para *Num* ocorre por um processo morfofonêmico, enquanto que o *Não* pós-verbal, sempre tônico, é uma categoria máxima XP<sup>5</sup> 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, como em (9). Santos, Araújo e Pereira (2018) também constataram que a Negação Pós-verbal é favorecida em contextos de

<sup>4</sup>Segundo Mioto, Silva e Lopes (2018, p. 54), a relação de dominância é definida como sendo: “ $\alpha$  domina  $\beta$  se e somente se existe uma sequência conexa de um ou mais galhos entre  $\alpha$  e  $\beta$  e o percurso de  $\alpha$  até  $\beta$  através dos galhos é unicamente descendente”.

<sup>5</sup>Conforme Mioto, Silva e Lopes (2018, p. 52), uma categoria XP é a projeção máxima de um sintagma.

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 distinção entre as três variantes da 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 — dita<sup>6</sup>, como em (13), ou inferível<sup>7</sup>, como em (14) — 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, em (13). Segundo Schwenter (2005, p. 1450), a restrição aparenta ser bastante rígida, de modo a ocorrer com a repetição do verbo da pergunta.

(13) Você gostou da festa da Ana?<sup>8</sup>

(Não) gostei não.

(14) 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 em que a Negação Pós-verbal é proferida quando a informação contextual não é dita, mas inferível. Em Freitag e Pinheiro (2020), a variável “estatuto informacional” foi controlada, mas sem significância, o que indica não ser um fator tão influente — cabe a ressalva de que Freitag e Pinheiro (2020) apenas controlaram as variantes “informação dada” e “informação nova”, sem menção a “informação inferível”.

Assim, percebe-se o quão complexo é o fenômeno da negação no português brasileiro, visto que demanda uma gama de perspectivas teóricas para dar conta de sua análise. Nesta pesquisa, ao considerar os diversos estudos sociolinguísticos, foi procedida uma meta-análise de modo a desvendar como são usadas cada variante quantitativamente, além de identificar quais restrições sociais são influentes no condicionamento da negação, um trabalho de vertente sociolinguística, como é apresentado na seção a seguir.

---

<sup>6</sup>Já ocorrido no texto anteriormente (Cunha, 2023, p. 166).

<sup>7</sup>Identificável mediante um processo de inferência (Cunha, 2023, p. 166).

<sup>8</sup>Os exemplos (13) e (14) foram construídos introspectivamente tomando como base os exemplos de Schwenter (2005).

### 3 FUNDAMENTAÇÃO TEÓRICA

Nesta seção, são descritas a área em que o presente estudo se baseia (Sociolinguística Variacionista), assim como a concepção de língua tomada como base, e as variáveis independentes sexo/gênero, escolaridade e faixa etária investigadas na meta-análise, que fazem parte do problema de restrição.

#### 3.1 Sociolinguística Variacionista

O presente trabalho parte da Teoria da Variação e Mudança ou Sociolinguística Variacionista (Labov, 2008), que propõe investigar a língua em seu contexto de uso. Para essa perspectiva teórica, “a língua é uma forma de comportamento social” (Labov, 2008, p. 215), logo exibe regularidades conforme a comunidade de fala<sup>9</sup> que a utiliza. Assim, essa perspectiva entende o sistema linguístico vinculado às pressões sociais a que os falantes estão submetidos, e não isolada desses contextos, como o fazem teorias formalistas, como o Estruturalismo e a Teoria Gerativa.

Nessa linha, percebe-se que a língua não apresenta apenas as regras categóricas — regras que se aplicam sempre da mesma maneira —, como as postuladas por modelos que consideram a linguagem em sua homogeneidade, mas também as regras ditas variáveis, que são as formas que podem se alternar na fala/escrita a depender de determinadas restrições (extra)linguísticas. Destarte, a Sociolinguística Variacionista se debruça sobre a heterogeneidade ordenada da língua ao observar que as variedades das comunidades não são caóticos, mas apresentam regularidades passíveis de estudo científico (Mollica; Braga, 2020).

Justamente por serem regulares, há algo guiando a variação entre as formas. Conforme Coelho *et al.* (2010, p. 96), foi proposto o “problema da restrição” para avaliar como fatores linguísticos e sociais se inter-relacionam para influenciar o falante a fazer uso de determinada forma. 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.

---

<sup>9</sup>Entendida aqui como “[...] um grupo que compartilha as mesmas normas a respeito da língua” (Labov, 2008, p. 188).

O fenômeno linguístico variável é codificado, na análise, como sendo uma variável dependente (VD). A VD, segundo Coelho *et al.* (2010, p. 131), pode ser de diversos tipos: binária, quando há apenas duas formas concorrendo; ternária, quando há três formas; ou eneária, quando há quatro ou mais variantes. No caso deste estudo, a VD é ternária, pois há três formas de negação verbal possíveis no PB.

A partir da variável dependente, definem-se as possíveis restrições, entendidas como variáveis independentes (VI). Essas variáveis são linguísticas, quando se referem a aspectos da estrutura gramatical (presença de determinante, paralelismo formal, tipo de oração etc.), ou sociais, quando se referem a aspectos da estrutura social em que o falante se situa (sexo/gênero, escolaridade, faixa etária etc.). Esta meta-análise, em especial, foca apenas na relevância das VIs sociais como mecanismo de restrição.

A influência das VIs sobre a VD é medida pelo método estatístico adotado pela Sociolinguística Variacionista, “[...] o *quantum* com que cada categoria postulada contribui para a realização de uma ou de outra variante das formas em competição” (Naro, 2020, p. 16, grifo do autor), que propõe investigar os dados obtidos de amostras da comunidade de fala com o uso de cálculos matemáticos.

Nesse processo, os modelos estatísticos proporcionam uma resposta abrangente sobre como a variação se dá na comunidade estudada, podendo se ter um panorama de como, de fato, as línguas humanas são usadas, distante das normatizações da gramática tradicional e das idealizações e introspecções de modelos formalistas de estudo linguístico. A Sociolinguística Variacionista, portanto, analisa a fala e a escrita cientificamente com *corpus* coletado pelo pesquisador na comunidade estudada, e os resultados são passíveis de comparação por apresentarem relativa padronização na metodologia.

## **3.2 Variáveis Sociais: Sexo/Gênero, Faixa Etária e Escolaridade**

### **3.2.1 Sexo/gênero**

A variável sexo/gênero é uma das mais usadas nos estudos sociolinguísticos para investigar as pressões sociais em fenômenos variáveis. Do ponto de vista social, homens e mulheres podem ocupar espaços diferentes na sociedade, participar de grupos diferentes e, desse modo, apresentar comportamentos distintos. Conforme Labov (2008, p. 348), “[...] a diferenciação sexual da fala frequentemente desempenha um papel importante no mecanismo da evolução linguística”.



Nesse sentido, os pesquisadores estudam se essa distinção também é levada para a comunidade de fala. Em diversos fenômenos do português, sexo/gênero é um fator estatisticamente significativo para explicar a variação de formas alternantes. Como mostra Paiva (2020), 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 a que pessoas do sexo masculino. De forma contrária, a autora também mostra que um estudo de William Labov 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 uma concordância entre alguns dos estudos sobre o papel dos sexos/gêneros na variação e na mudança linguísticas — até porque sociedades diversas têm comportamentos diversos.

Outra pesquisa acerca dessa variável é a de Scherre e Yacovenco (2011), que, utilizando dados de pesquisas sobre a variação entre as formas “tu” e “você”, apresenta uma dicotomia no papel das mulheres, que ora utilizam mais “tu” do que homens (em dados do Sul e do Norte), ora utilizam menos (em dados do Rio de Janeiro). Com esses dados, as autoras associam o uso maior de “tu” pelas mulheres quando essa variante se torna um traço da identidade regional (no Sul, parte do Nordeste e Norte) e o uso menor quando não há esse traço, mas apenas uma relação de intimidade entre os falantes (no Rio). Nesse contexto, além de também se basearem em pesquisas sobre a variação no imperativo e na concordância verbal, Scherre e Yacovenco (2011) chegam à generalização de que mulheres estão à frente em processos de mudança envolvendo variantes menos marcadas<sup>10</sup> (não necessariamente mais prestigiadas), enquanto homens estariam à frente em variantes mais marcadas (não necessariamente menos prestigiadas).

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, V., 2020, p. 71). Como exemplo, há a revisão baseada em meta-análise realizada por Novais e Siqueira (2020) em seis trabalhos sociolinguísticos alagoanos realizados com uma amostra coletada na região sertaneja (Vitório, 2020), realizada com o objetivo de analisar a influência da variável sexo/gênero no condicionamento dos fenômenos. O artigo reporta que apenas um trabalho apresentou significância dessa variável, 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

---

<sup>10</sup>Entende-se “marcação” no texto como sendo formas mais/menos dependentes das relações interacionais ou mais/menos frequentes ou mais/menos aceitos socialmente (Scherre; Yacovenco, 2011, p. 138-139).

independente. Logo, conforme Paiva (2020) pontua, as transformações na sociedade podem levar a uma neutralização ou acentuação da variável sexo/gênero.

Recentemente, a maneira de codificar essa variável vem sendo problematizada por pesquisadores. Freitag (2015) analisa os problemas acerca do modo como se concebem as ideias de sexo e de gênero, além de relacioná-las ao comportamento linguístico do falante. Nas conclusões da autora, parece ter suporte a adoção dos modelos de comunidade de prática, em que não se analisa a variação de forma macro, mas, sim, micro (contextos mais específicos). Portanto, mesmo que ocorra alteração de padrões sociais relativos ao gênero do falante, o que é comum nas sociedades, os sociolinguistas não ficariam reféns de uma variável e uma hipótese clássica — a de que o sexo/gênero do falante interfere no seu comportamento linguístico —, que pode falhar, como aborda Freitag (2015).

### 3.2.2 Faixa etária

Um dos processos estudados dentro da Sociolinguística Variacionista é a mudança linguística. Segundo a Sociolinguística, a língua sofre pressões estruturais e sociais ao longo do tempo, que levam a alterações no uso desta pela comunidade de fala. Essas alterações podem acarretar tanto um processo de variação estável, quando as formas convivem harmoniosamente, ou de mudança, quando uma variante suplanta outra dentro do contexto linguístico em que ocorrem.

No estudo da mudança da língua, a variável faixa etária pode demonstrar se há ou não um processo de mudança linguística ocorrendo, pois, como é estratificada em faixas que englobam jovens, adultos e idosos, ela possibilita visualizar como falantes que viveram em diferentes épocas utilizam determinada forma linguística. 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 de fala, pois a fala de uma pessoa reflete o seu processo de aquisição da linguagem, em uma das hipóteses sobre o processo de mudança linguística. Considerando um caso hipotético, um senhor de 60 anos apresenta mais conservadorismo no uso linguístico porque o seu processo de aquisição do vernáculo ocorreu há mais de 40 anos, logo ele pode refletir um estado anterior da língua. Esse tipo de análise dos processos de mudança é chamado de Mudança em Tempo Aparente (Naro, 2020, p. 45).

Nas pesquisas atuais, essa variável é geralmente estratificada em diferentes faixas etárias, que recebem rótulos como F1, F2 e F3, compreendendo de falantes recém adultos (+18) até pessoas com mais idade (44 ou mais), como é o caso da amostra coletada para o projeto A

*Língua Usada no Sertão Alagoano* (Vitório, 2020). Entretanto, essa não é a única maneira, pois também a faixa etária pode ser medida numericamente pela quantidade de anos ou pelas categorias de “jovem”, “adulto” e “idoso”. De todo modo, a faixa etária é uma importante pressão social estudada dentro da Sociolinguística Variacionista.

Freitag (2005) aponta que esse fator não engloba somente idades, mas sim um feixe de outras variáveis envolvidas com ela, como as interações sociais — um adulto, por exemplo, pode receber mais imposições linguísticas normativas do que um adolescente, pois aquele frequentaria, em geral, ambientes como o trabalho, local onde é exigida formalidade, em alguns casos. Além disso, outro problema a se pensar é a própria segmentação das faixas, visto que não há consenso entre os teóricos. Devido a isso, as pesquisas sociolinguísticas podem apresentar distribuições diferentes, seja tomando como F1 pessoas de 18-25 anos ou mesmo de 15-18 (o mesmo ocorre em outras faixas etárias). Nesse sentido, encontrar paridade entre os trabalhos é mais difícil porque há grupos de pessoas distintos sendo agrupados na mesma categoria, o que pode restringir o poder de generalização da meta-análise aqui empreendida.

### 3.2.3 A variável escolaridade

A escola, nas sociedades ocidentais, desempenha um importante papel no desenvolvimento da competência comunicativa nos estudantes, ou seja, a capacidade de usar a língua em diferentes contextos sociais (Bortoni-Ricardo, 2004, p. 73). Entretanto, 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*<sup>11</sup> (Cavalcanti, 2013).

Com isso, a escola 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 rotacismo, por exemplo, que é a troca do /l/ pelo /r/, é julgado socialmente como

---

<sup>11</sup>Este foi um livro didático escrito por Heloísa Ramos distribuído pelo MEC à Educação de Jovens e Adultos (EJA), criticado por inúmeros veículos de imprensa por colocar como “adequadas” construções do tipo: “nós pega o peixe”. Entre os comentários negativos, ressoava o discurso de que o MEC promovia a ignorância e os erros de português por aprovar o livro.

vindo de um indivíduo pobre 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 a variável escolaridade que investigam traços graduais, isto é, formas que perpassam todas as camadas sociais, tendem a encontrar pouca ou nenhuma significância estatística.

Mesmo assim, as pesquisas sociolinguísticas do Brasil recorrem a tal pressão social por ser uma variável clássica. Em geral, a escolaridade é estratificada em E0 (analfabetos), E1 (ensino fundamental), E2 (ensino médio) e E3 (ensino superior) — embora haja estratificações com o número de anos de escolaridade e com outras conformações da educação básica (ginásio, colegial). Os dois extremos da variável, E0 e E3, são interessantes de serem contrastados entre si porque o ensino superior torna o falante da língua, para os padrões da sociedade, culto<sup>12</sup>(Faraco, 2002), enquanto a falta de alfabetização incute no indivíduo processos de forte estigmatização social e linguística. Felizmente, entretanto, o número de analfabetos no Brasil vem diminuindo com a nacionalização da educação formal (Tokarnia, 2020), o que favorece o processo de ascensão social e acaba levando, algumas vezes, ao desuso de formas estigmatizadas.

Para a Sociolinguística, então, vem sendo cada vez mais difícil preencher essa célula social, o que leva pesquisadores a focarem no contraste entre a feitura ou não do curso superior, que ainda é um mecanismo de distinção social por meio da língua. Nesse sentido, não há uma regularidade na estratificação da escolaridade em todos os estudos científicos que analisam a língua em uso, devendo, então, se tomar cuidado na generalização dos dados em revisões de literatura.

---

<sup>12</sup>Nos termos do NURC, considera-se culto um falante que tem ensino superior completo.

## 4 METODOLOGIA

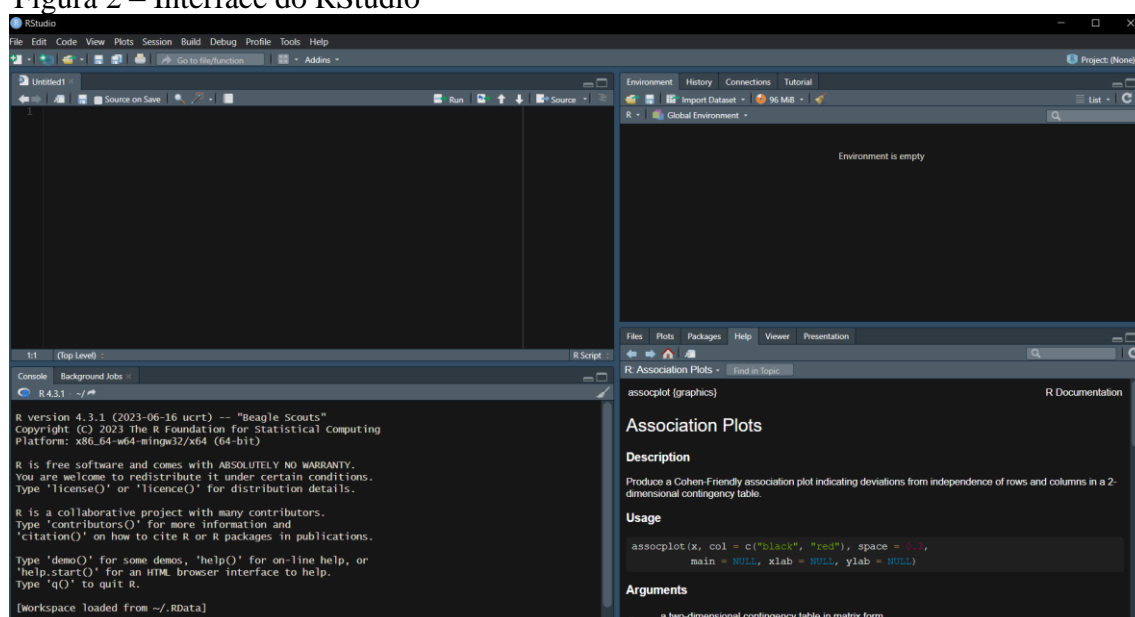
Nesta seção, são descritas as ferramentas estatísticas utilizadas e como funciona a meta-análise. Assim, este capítulo permite entender como a pesquisa foi feita para que outras meta-análises possam se basear para a sua realização.

### 4.1 Ferramentas Estatísticas Utilizadas: R e RStudio

O R é uma linguagem de programação utilizada em muitas pesquisas estatísticas pelo seu maquinário desenvolvido para análises de dados. Recentemente, dentro do campo da Linguística, pesquisadores vêm defendendo o seu uso por apresentar mais ferramentas e ser mais maleável do que programas clássicos usados em Sociolinguística, como o GoldVarb X (Godoy, 2019; Gries, 2019; Oushiro, 2017). Outro aspecto que torna o R mais útil é o fato de ser usado em diversas áreas científicas, enquanto o GoldVarb X é mais restrito à Sociolinguística, logo se torna mais útil utilizar ferramentas que sejam melhor entendidas por outros pesquisadores.

Ao se trabalhar com o R, pode se fazer uso da interface RStudio (R Core Team, 2022) por ser mais intuitiva. Com ela, pode-se criar *scripts*, isto é, códigos de programação com as funções que o pesquisador deseja que o RStudio execute. A Figura 2 apresenta a interface desse programa computacional.

Figura 2 – Interface do RStudio



Fonte: elaborada pelo autor (2023).

Ademais, no RStudio, o pesquisador tem a liberdade de executar diversos pacotes de funções estatísticas desenvolvidas por outros pesquisadores. Isso só é possível porque o R é um *software* de código aberto, o que favorece a solução de problemas pela própria comunidade usuária do programa, além de promover a criação de novas ferramentas estatísticas para pesquisas quantitativas.

Para esta pesquisa especificamente, foram utilizados pacotes de funções específicas. O pacote *ggstatsplot* (Patil, 2021) foi utilizado para criar gráficos, o pacote *sjPlot* (Lüdtke, 2023) foi usado para criação das tabelas com modelos de regressão, o pacote *tidyverse* (Wickham *et al.*, 2019) foi utilizado para utilizar ferramentas de filtragem de dados e o pacote *nnet* (Venables; Ripley, 2002) foi usado para a feitura das regressões multinomiais. Por meio dessas ferramentas, foi possível realizar a meta-análise, embora esta também pudesse ser feita por meio de outros pacotes, como o *metafor* (Viechtbauer, 2010)<sup>13</sup>.

## 4.2 Meta-análise

De modo a investigar estatisticamente a relevância dos tradicionais condicionamentos sociais (sexo/gênero, escolaridade e faixa etária), foi usada a meta-análise (Freitag, 2021), uma técnica de análise de revisões que agrupa os dados das pesquisas e aplica a eles testes estatísticos comuns 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 já 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 na sua adequada realização: 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 funcionamento do fenômeno. A partir daí, deve-se fazer a busca sistematizada dos

---

<sup>13</sup>A título de curiosidade, quem indicou o uso desse pacote foi o *ChatGPT*, a nova Inteligência Artificial treinada pela OpenAI com ampla quantidade de textos disponíveis na internet até 2021 de modo a gerar textos bem estruturados em diversas línguas. A partir do comando “Qual pacote do R serve para fazer meta-análise?”, a IA disse que o *metafor* é o pacote mais utilizado.

trabalhos, controlando o tempo (período de publicação dos artigos) 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 artigos organizados e filtrados, passa-se à análise estatística destes, seja por meio de comparação gráfica, modelos estatísticos de ajuste de dados ou análise da distribuição dos resíduos dos estudos com testes de qui-quadrado. 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, logo seu uso deve ser incentivado dentro da Sociolinguística Variacionista.

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) com pesquisas da Bahia, que reporta que a fala culta de Salvador e Feira de Santana favorece a concordância padrão, enquanto variedades populares e afro-brasileiras desfavorecem-na.

Apesar de incipiente, esse método de pesquisa mostra-se frutífero por agrupar os diversos estudos sobre os fenômenos linguísticos variáveis e proporcionar generalizações mais confiáveis acerca da variação e da mudança linguística. Nesse sentido, a meta-análise é uma ferramenta proveitosa de ser utilizada e, por isso, é o método escolhido no presente trabalho.

#### 4.3 Medidas Estatísticas

Nesta meta-análise, alguns termos e testes e estatísticas foram realizados, logo, necessitam da sua devida explanação. O primeiro deles é o valor de  $p$ , convencionalmente utilizado como valor crítico em 5%, um valor de significância que diz respeito à probabilidade de um evento ter ocorrido ao acaso (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 (descarta-se a possibilidade de o resultado ser ao acaso).

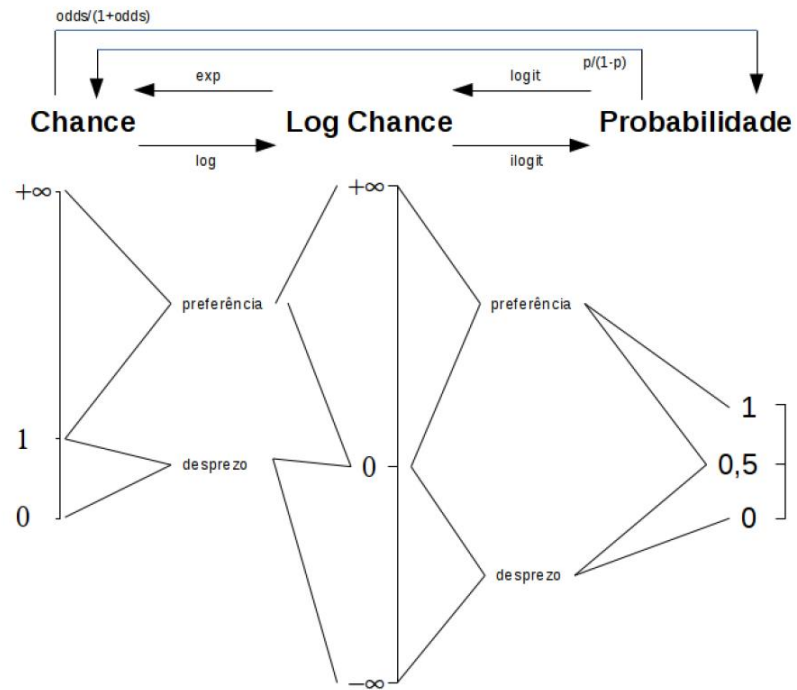
Entretanto, o valor de  $p$  é sensível ao tamanho amostral, por isso, pesquisas que tendem a comparar a significância de variáveis podem ter problemas se as amostras forem de tamanhos distintos. Resolver esse problema implica utilizar uma medida conhecida como tamanho de efeito, que analisa de modo 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” a “Muito forte”. 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, na população, há 95% de certeza de o parâmetro se encontrar no intervalo — de modo mais técnico, 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 — o intervalo varia de tamanho para garantir os 95% de certeza.

Outro método utilizado nesta pesquisa é a regressão logística — e a regressão multinomial, que é um tipo de regressão logística para casos em que a variável dependente tem mais de dois níveis. 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. 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). Na figura 3, explicita-se por que o valor de *log-odds* é preferido na análise: valores abaixo de 0 indicam desfavorecimento; acima de 0, favorecimento (Oushiro, 2017, p. 186).



Figura 3 – Comparação entre Chances, Log-odds (Log Chance) e Probabilidade



Fonte: Gries (2019, p. 265).

## 5 RESULTADOS

Nesta seção, são apresentados os resultados da meta-análise empreendida. De início, é apresentado o resultado da revisão sistemática e, após isso, uma análise geral das pesquisas é reportada. Por fim, em três seções distintas, apresentam-se os resultados para cada condicionamento social selecionado, nesta ordem: sexo/gênero, escolaridade e faixa etária.

### 5.1 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 títulos<sup>14</sup>, textos que tivessem a palavra “negação” e que pudessem ter pelo menos uma das seguintes palavras: Português, português<sup>15</sup>, 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 artigos 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. São eles: Braga e Silva (2011); Pessôa (1985); Mendes (1996); Rocha (2012); Júnior (2011); Junior (2011); Araújo (2012); Rocha (2013); Souza (2017); Marinho (2019); Namiuti (2008); Sousa (2018a); Belletti (2020); Silva e Mello (2015); Meireles (2001); Namiuti e Miotto (2014); Nascimento (2014); Gonçalves, Temponi e Pacheco (2014); Castro *et al.* (2017); Pinto (2015); Reimann e Yacovenco (2011); Reimann e Yacovenco (2014); Nunes (2014); Yacovenco e Nascimento (2016); Sousa (2018b); Oliveira, Carvalho e Silva e Prata (2007); Namiuti-Temponi (2009); Sousa (2012); Silva (2016); Santos (2021); Alencar (2009); Santos Silva (2009); Serra (2018); Sousa (2007); Lopes da Silva (2020); Pinto (2017); Lima [20--?]; Graça e Avelar (2020); Cavalcante (2018); Silveira (2004); Cavalcante (2007); Lamberti Nunes

---

<sup>14</sup> 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.

<sup>15</sup> Alguns trabalhos fazem uso da palavra com inicial minúscula, então foram colocadas as duas formas.

(2013); Marcelino (2018); Namiuti (2007); e Moreno (2005). Vale ressaltar que o trabalho de Mendes (1981)<sup>16</sup> não pôde ser incluído porque o *site* tornou o acesso a este restrito.

Após isso, procedeu-se uma análise dos resumos dos artigos para filtrar ainda mais os textos com os mesmos critérios anteriores, além de se acrescentar o critério de o arquivo conter o texto inteiro, e não apenas um pedaço, como foi o caso de Lamberti Nunes (2013), Pinto (2017) e Lima [20--?]. 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).

O passo final da filtragem de artigos foi analisar quais disponibilizam os dados completos da pesquisa para seu uso na meta-análise, além de conterem amostras de dados que sejam significativamente numerosas (não contenham dados de apenas um ou dois informantes<sup>17</sup>). 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, Reimann e Yacovenco (2011) foi descartado na análise dos condicionamentos sociais por amalgamar os dados da Dupla Negação com a Negação Pós-verbal; também Reimann e Yacovenco (2014), Rocha (2012), Serra (2018) e Yacovenco 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). 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).

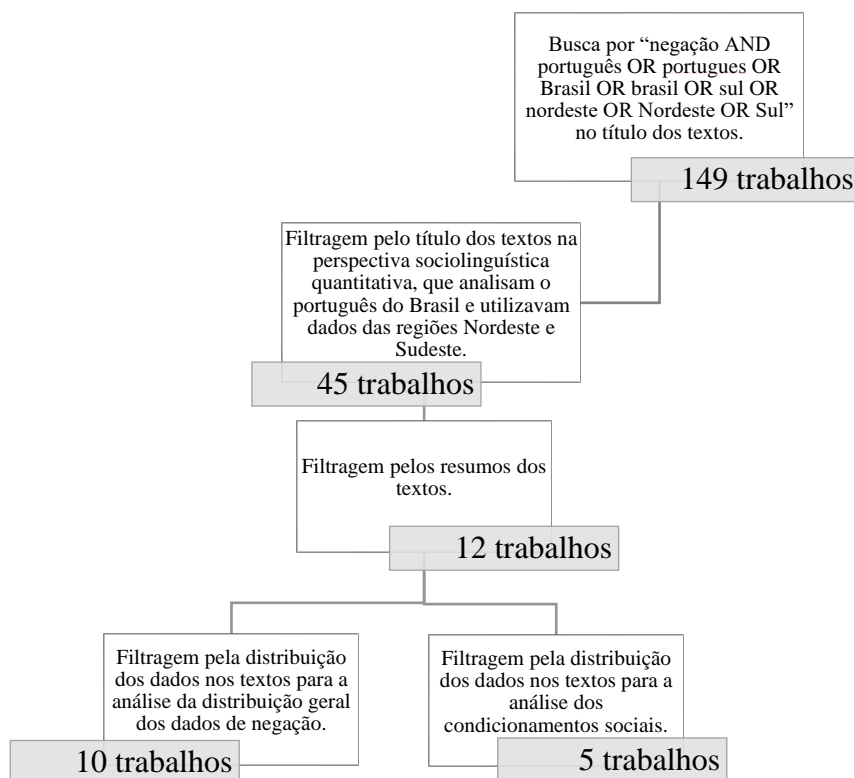
O fluxograma a seguir indica graficamente como foi feita a filtragem dos trabalhos.

---

<sup>16</sup>MENDES, Ana Catarina Fabrício. **Uma contribuição ao estudo da negação em português**. 1981. Dissertação (Mestrado em Língua Portuguesa) – Pontifícia Universidade Católica de São Paulo, São Paulo, 1981. Disponível em: <https://tede2.pucsp.br/handle/handle/14174>. Acessado em: 28 jun. 2022.

<sup>17</sup>Para evitar uma diferença muito grande de informantes entre os trabalhos.

### Fluxograma – Revisão sistemática de literatura



Fonte: elaborado pelo autor (2023).

## 5.2 Análise Geral

Das 10 pesquisas selecionadas para análise, 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 numa comunidade afrodescendente baiana, o que sustentaria uma hipótese de essa negação ter sido introduzida no português brasileiro pelo contato linguístico entre o Brasil e países africanos durante a escravidão.

Tabela 1 – Distribuição da negação nas pesquisas selecionadas pela revisão sistemática

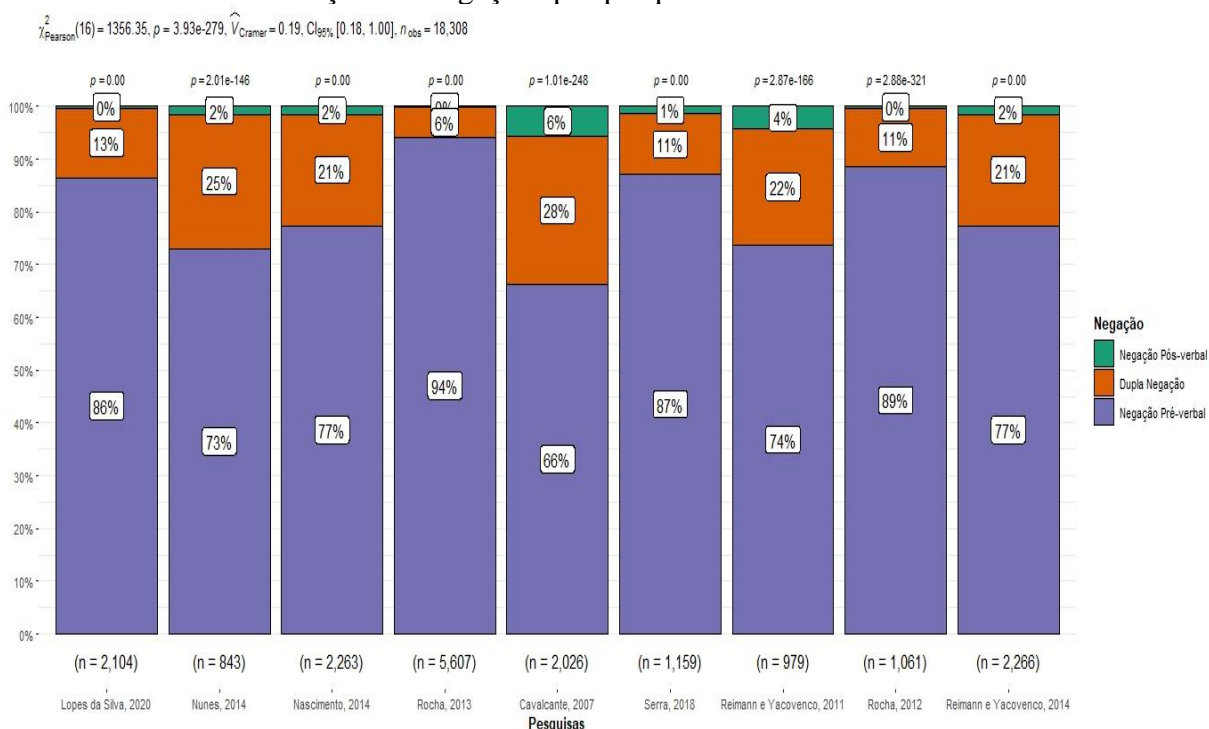
| Pesquisa                      | Região   | Estado         | Amostra       | Infor<br>mant<br>es | Pré-verbal      | Dupla<br>Negação | Negação<br>Pós-verbal | Total          |
|-------------------------------|----------|----------------|---------------|---------------------|-----------------|------------------|-----------------------|----------------|
| Lopes da Silva2020            | Nordeste | Ceará          | Torres (2008) | 40                  | 1819<br>(86,7%) | 278<br>(13,2%)   | 7 (0,3%)              | 2097<br>(100%) |
| Nunes2014                     | Sudeste  | Rio de Janeiro | PEUL          | 8                   | 616 (73%)       | 214<br>(25,3%)   | 13 (1,5%)             | 843<br>(100%)  |
| Nascimento2014                | Sudeste  | Espírito Santo | PortVix       | 18                  | 1751<br>(77,3%) | 478<br>(21,1%)   | 34 (1,5%)             | 2263<br>(100%) |
| Rocha2013                     | Sudeste  | São Paulo      | GESOL         | 48                  | 5279<br>(94,1%) | 324<br>(5,7%)    | 4 (0,07%)             | 5607<br>(100%) |
| Cavalcante2007                | Nordeste | Bahia          | Vertentes     | 18                  | 1343<br>(66,2%) | 568<br>(28%)     | 115 (5,6%)            | 2026<br>(100%) |
| Serra2018                     | Nordeste | Maranhão       | Própria       | 24                  | 1009<br>(87,5%) | 133<br>(11,4%)   | 17 (1,4%)             | 1159<br>(100%) |
| Reimmann e<br>Yacovenço2011   | Sudeste  | Espírito Santo | PortVix       | 8                   | 721<br>(73,6%)  | 216<br>(22%)     | 42 (4,2%)             | 979<br>(100%)  |
| Rocha2012                     | Sudeste  | São Paulo      | GESOL         | 12                  | 940<br>(88,6%)  | 117<br>(11%)     | 4 (0,4%)              | 1061<br>(100%) |
| Yacovenço e<br>Nascimento2016 | Sudeste  | Espírito Santo | PortVix       | 18                  | 1751<br>(77,3%) | 478<br>(21,1%)   | 34 (1,5%)             | 2263<br>(100%) |
| Reimmann e<br>Yacovenço2014   | Sudeste  | Espírito Santo | PortVix       | 18                  | 1754<br>(77,4%) | 478<br>(21,1%)   | 34 (1,5%)             | 2266<br>(100%) |

Fonte: elaborada pelo autor (2023).

Outro aspecto a se notar nessa comparação é a semelhança entre as pesquisas de Yacovenço e Nascimento (2016) e Reimann e Yacovenço (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; Yacovenço, 2014) e, depois, publicaram em um periódico (Yacovenço; Nascimento, 2016). Para visualização e teste de hipótese, o gráfico 1 apresenta os resultados sem considerar Yacovenço e Nascimento (2016)<sup>18</sup>.

<sup>18</sup> No apêndice A está contido o *script* utilizado na análise geral.

Gráfico 1 – Distribuição das negações por pesquisa selecionada na revisão sistemática



Fonte: elaborado pelo autor (2023).

Como foi observado no gráfico, a diferença entre os 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. Já sobre a Negação Pós-verbal, Cavalcante (2007) e Reimann e Yacovenco (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 forma negativa. Assim, a hipótese, presente na seção de introdução, da Negação Pós-verbal e Dupla Negação serem mais favorecidas no Nordeste em razão do contato linguístico fica desfavorecida.

### 5.3 Análise da Variável Sexo/gênero

Na análise da variável “sexo/gênero”<sup>19</sup>, 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

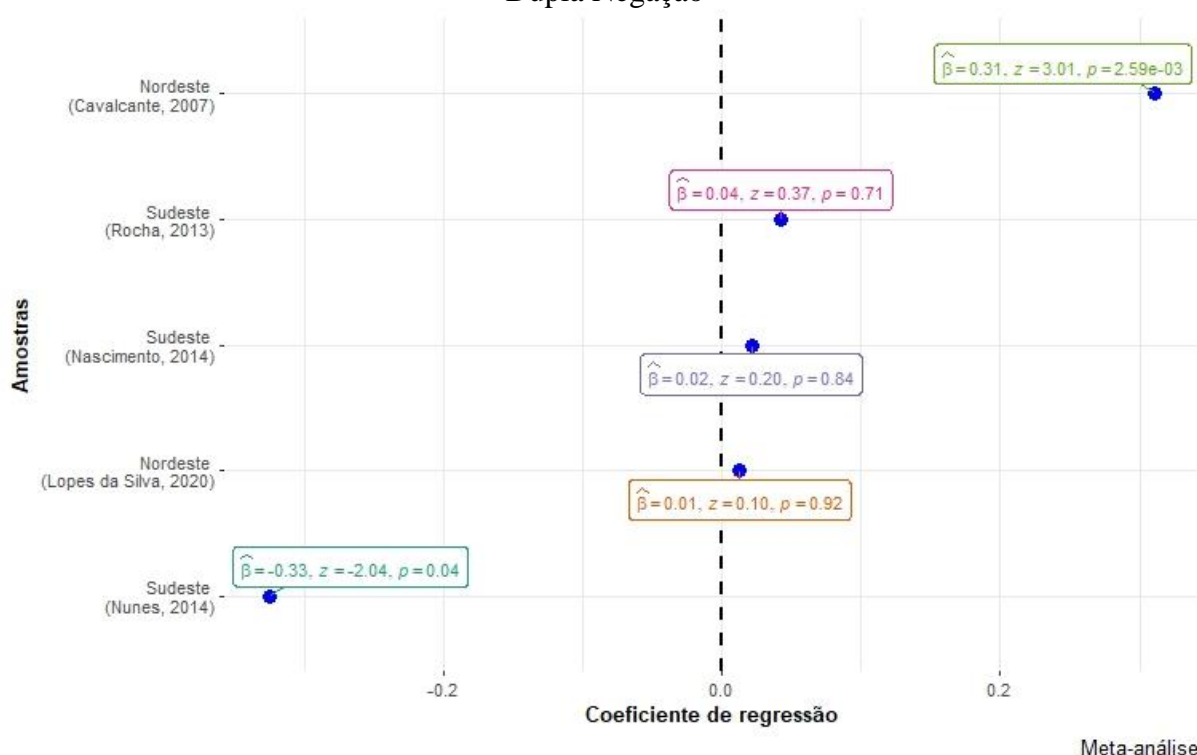
<sup>19</sup> Script disponível no apêndice B.

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 Negação<sup>20</sup>, que utilizou três trabalhos. Assim, foi possível analisar como cada negação variava em relação ao nível de referência “Negação Pré-verbal”.

A influência da variável “sexo/gênero” no condicionamento da Dupla Negação está exposta no gráfico 2 a seguir. Nele, foi utilizado como valor  $\beta$  o quanto variava em *log-odds* pela alteração do valor “masculino” para “feminino”, pois se tem uma ideia da influência de tal variável social no uso da negação.

Gráfico 2 – Influência da variável “sexo/gênero” na variação entre Negação Pré-verbal e Dupla Negação



Fonte: elaborado pelo autor (2023).

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,

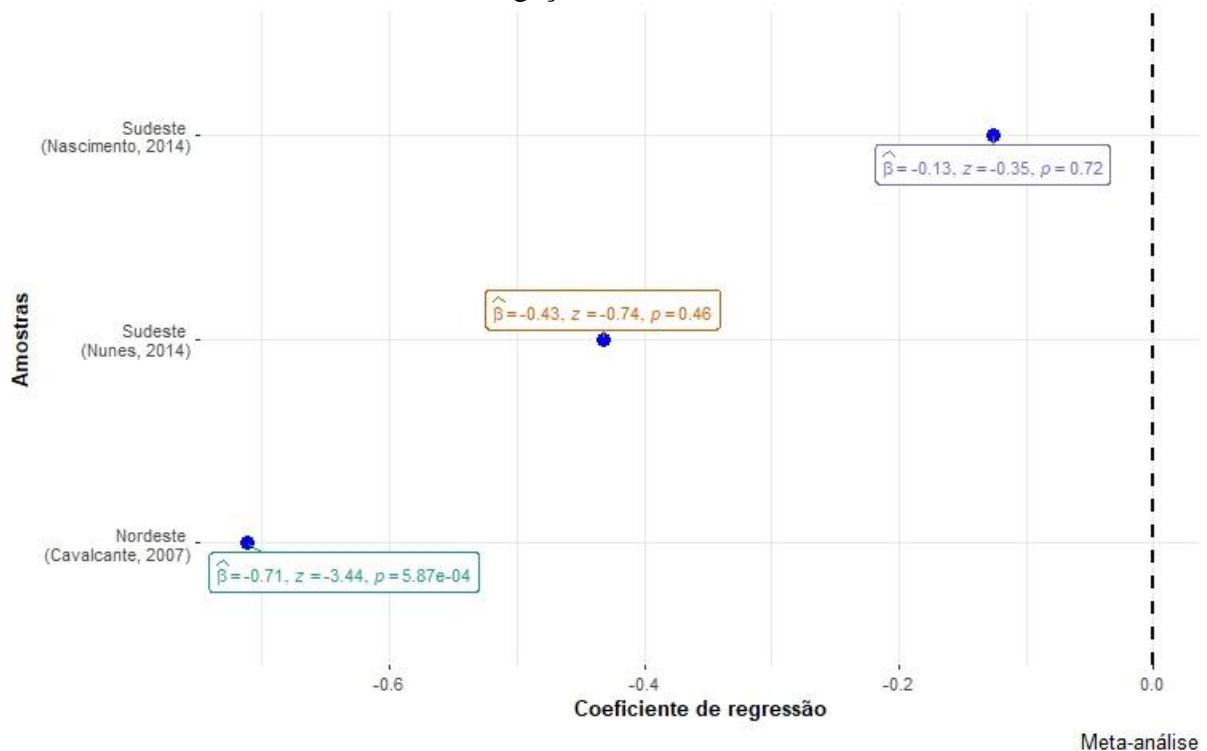
<sup>20</sup>Foi escolhida essa relação, e não a entre Negação Pré-verbal e Negação Pós-verbal, porque as duas são utilizadas em contextos diferentes, conforme apontou Schwenter (2005).

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 variação significativa — mesmo Nunes (2014) tem um valor de  $p$  muito próximo do limiar convencionado 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 da 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  $\beta$  o quanto variava em *log-odds* pela alteração do nível da variável.

Gráfico 3 – Influência da variável “sexo/gênero” na variação entre a Dupla Negação e a Negação Pós-verbal



Fonte: elaborado pelo autor (2023).



Nos três trabalhos que apresentavam dados para a Negação Pós-verbal em cada sexo/gênero, nota-se a mesma realidade. O único trabalho significativo é o de Cavalcante (2007), que encontrou um número alto de dados de Negação Pós-verbal. Nos outros trabalhos, entretanto, a influência da variável foi baixa ou nula. Nos dois casos, como se viu, a hipótese alternativa foi aceita, porque não houve influência mais geral do sexo/gênero no uso da negação. Devido a isso, é cabível pensar que novos *corpora* desenvolvidos para analisar a negação verbal podem prescindir da segmentação das células sociais em sexo/gênero, visto que, independentemente de ser homem ou mulher, o informante utilizaria a negação da mesma maneira, seja no Nordeste ou no Sudeste.

#### 5.4 Análise da Variável Escolaridade

Na análise da variável “escolaridade”<sup>21</sup>, 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.

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:

Tabela 2 – Influência da variável “escolaridade” em Cavalcante (2007)

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

Fonte: elaborada pelo autor (2023).

<sup>21</sup>Script disponível no apêndice C.

Curiosamente, essa variável não foi significativa na análise do autor, o que pode se dar por a pesquisa dele 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.

A nível de probabilidade, ficam mais evidentes os resultados. O uso da Dupla Negação quando a escolaridade do falante é “semialfabetizado” é de -1,07 *log-odds*, o que equivale a uma probabilidade de 25,5% de uma forma dessa aparecer no discurso de alguém com mais escolarização<sup>22</sup>. 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. Já 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, mas o da Negação Pós-verbal não tanto, o que influencia o valor baixo de  $R^2$ .

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.

Tabela 3 – Influência da variável “escolaridade” em Lopes da Silva (2020)

| <i>Predictors</i>              | <b>VD</b>       |               |          |
|--------------------------------|-----------------|---------------|----------|
|                                | <i>Log-Odds</i> | <i>CI</i>     | <i>p</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           |               |          |

Fonte: elaborada pelo autor (2023).

<sup>22</sup>Transformação de *log-odds* para probabilidade feita a partir da função *ilogit()*. Conferir Oushiro (2017) para mais detalhes.

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 indica 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.

Tabela 4 – Influência da variável “escolaridade” em Rocha (2013)

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

Fonte: elaborada pelo autor (2023).

Como se observa, a variável (nessa pesquisa com os níveis “ensino médio” e “ensino superior”) 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.

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. Nesse sentido, pode-se presumir que se trataria de sujeitos que estudaram o Ensino Fundamental e o Médio.

Tabela 5 – Influência da variável “escolaridade” em Nunes (2014)

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

Fonte: elaborada pelo autor (2023).

Observa-se na tabela que a escolarização só foi influente 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 sujeito a usar a Dupla Negação no Rio de Janeiro.

Por fim, a tabela 6 apresenta a regressão multinomial criada para dar conta dos dados de Nascimento (2014). Nesse estudo, a escolaridade foi elaborada com os níveis “Ensino Fundamental”, “Ensino Médio” e “Ensino Universitário”, ou seja, abrange os três níveis de escolarização comuns no Brasil. Nesse sentido, há uma análise mais abrangente da influência dessa variável.

Tabela 6 – Influência da variável “escolaridade” em Nascimento (2014)

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

Fonte: elaborada pelo autor (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. 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) e talvez sofram estigmatização.

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 idiossincrasia de São Paulo, a princípio. Nesse sentido, é possível concluir 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, visto que é uma fase em que as crianças estão começando a adolescência e são mais sensíveis às pressões sociais sobre seus comportamentos. 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.

## 5.5 Análise da Variável Faixa Etária

Para a análise da variável “faixa etária”<sup>23</sup>, pensou-se inicialmente que seria possível fazer um trabalho semelhante ao realizado para a variável “sexo/gênero”, mas não foi o caso. Mais uma vez, a variável foi estruturada de modo diferente em cada pesquisa, seja mudando os grupos por idade ou tomando duas faixas etárias em vez de três — o usual na Sociolinguística Variacionista. Por conta disso, é feita a análise de modo individual com uma síntese ao fim.

---

<sup>23</sup>Script disponível no apêndice D.

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.

Tabela 7 – Influência da variável “faixa etária” em Cavalcante (2007)

| <i>Predictors</i>                        | <i>Log-Odds</i> | <i>CI</i>     | <b>VD</b>        |  | <i>Response</i>    |
|--|-----------------|---------------|------------------|--|--------------------|
|  |                 |               | <i>p</i>         |  |                    |
| (Intercept)                              | -0.81           | -0.99 – -0.64 | <b>&lt;0.001</b> |  | 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 | <b>&lt;0.001</b> |  | Negação Pós-verbal |
| faixa.etaria41-60 anos                   | -1.02           | -1.54 – -0.49 | <b>&lt;0.001</b> |  | Negação Pós-verbal |
| faixa etaria [+60 anos]                  | 0.13            | -0.30 – 0.56  | 0.556            |  | Negação Pós-verbal |
| Observations                             | 2026            |               |                  |  |                    |
| R <sup>2</sup> / R <sup>2</sup> adjusted | 0.009 / 0.008   |               |                  |  |                    |

Fonte: elaborada pelo autor (2023).

Nesse trabalho, percebe-se que 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 R<sup>2</sup> baixíssimo (0,009).

Já, no estudo de Lopes da Silva (2020), a faixa etária foi binária: “<30 anos” e “>30 anos”. Por conta disso, a comparação direta com outros trabalhos fica prejudicada, limitando o poder explanatório da meta-análise. De todo modo, os resultados são expostos na tabela 8 abaixo.

Tabela 8 – Influência da variável “faixa etária” em Lopes da Silva (2020)

| <i>Predictors</i>       | <i>Log-Odds</i> | <b>VD</b>     |  | <i>p</i>         |
|-------------------------|-----------------|---------------|--|------------------|
|                         |                 | <i>CI</i>     |  |                  |
| (Intercept)             | -2.04           | -2.23 – -1.86 |  | <b>&lt;0.001</b> |
| faixa etaria [+30 anos] | 0.31            | 0.06 – 0.57   |  | <b>0.016</b>     |
| Observations            | 2097            |               |  |                  |
| R <sup>2</sup> Tjur     | 0.003           |               |  |                  |

Fonte: elaborada pelo autor (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 R<sup>2</sup> (0,003). Nesse sentido, percebe-se que, 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.

Tabela 9 – Influência da variável “faixa etária” em Rocha (2013)

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

Fonte: elaborada pelo autor (2023).

Nessa pesquisa, percebe-se que 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  $R^2$  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.

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. A falta de isonomia nos grupos já inviabilizaria a análise estatística, mas ela foi feita aqui. Os resultados da pesquisa são apresentados na tabela 10.



Tabela 10 – Influência da variável “faixa etária” em Nunes (2014)

| <i>Predictors</i>                        | <i>Log-Odds</i> | <i>CI</i>     | <b>VD</b> |                    |
|--|-----------------|---------------|-----------|--------------------|
|  |                 |               | <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 <sup>2</sup> / R <sup>2</sup> adjusted | 0.017 / 0.015   |               |           |                    |

Fonte: elaborada pelo autor (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 R<sup>2</sup> 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. Acerca dos resultados, eles estão dispostos na tabela 11 abaixo.

Tabela 11 – Influência da variável “faixa etária” em Nascimento (2014)

| <i>Predictors</i>                        | <i>Log-Odds</i> | <i>CI</i>     | <b>VD</b> |                    |
|--|-----------------|---------------|-----------|--------------------|
|  |                 |               | <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 <sup>2</sup> / R <sup>2</sup> adjusted | 0.001 / 0.001   |               |           |                    |

Fonte: elaborada pelo autor (2023).

Nesse estudo, a VI “faixa etária” não foi significativa em nenhum dos casos, 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 (embora, como se saiba, o intervalo de confiança fique próximo do zero). 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 — sempre observando a ressalva do intervalo de confiança estar próximo ao zero. 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 PROBLEMAS EM META-ANÁLISES DENTRO DA SOCIOLINGÜÍSTICA VARIACIONISTA

A realização da presente meta-análise foi limitada por diversos problemas relatados ao longo da apresentação dos resultados. Seja por as variáveis apresentarem diferentes segmentações ou por os trabalhos reportarem os dados da VD de modo binário ou ternário, seja pelos diferentes grupos em cada variável (diferença de intervalo de idade entre as pesquisas), foram diversas as razões para a comparação não ser mais bem executada. Nesse sentido, cabe ressaltar com mais clareza os problemas encontrados e apresentar algumas sugestões para o futuro da meta-análise dentro do campo da Sociolinguística Variacionista.

De início, destaca-se a dificuldade em encontrar os textos na revisão de literatura. Como as plataformas utilizam palavras-chave para busca de trabalhos, é possível que alguns estudos tenham ficado de fora por seus títulos não conterem os descritores-chave como “negação”, “variação”, “sociolinguística” e afins. Um exemplo disso é o trabalho de Freitag e Pinheiro (2020), cujo título é *Modelo de árvore de inferência condicional para explicar usos linguísticos variáveis*, que analisa a negação no estado de Sergipe por meio de árvores de inferência condicional. Como o trabalho não possui “negação” como título, não foi incluído na revisão sistemática, pois o Google Acadêmico não o identificou<sup>24</sup>.

O segundo ponto a ser mencionado é que não há uma padronização no modo como se definem as VIs. Com exceção da variável “sexo/gênero”, que, por convenção, vem sendo binária, com os níveis “masculino” e “feminino” em todos os trabalhos que a utilizam, as outras variáveis sociais, ou eram construídas de modo distinto e não comparável ou tinham níveis com intervalos diferentes (ora a faixa etária tinha um grupo com intervalo de 10 anos, ora com 20, ora com 19 anos). No caso da escolaridade, mais especificamente, ora se utilizava a quantidade de anos de escolarização, ora se utilizava uma definição confusa — como “semialfabetizado” em Cavalcante (2007)<sup>25</sup> —, ora se utilizava a quantidade de anos de escolarização em vez das etapas da educação básica. Toda essa assimetria dificulta a meta-análise e até impossibilitou a comparação dos resultados em um único gráfico semelhante ao feito para a variável “sexo/gênero”.

<sup>24</sup>Preferiu-se não adicionar textos que estivessem fora da pesquisa de revisão sistemática para manter a reprodutibilidade do processo.

<sup>25</sup>Na definição do autor, seriam semialfabetizados os indivíduos que tiveram algum tipo de contato com a escolarização, ou seja, poderia ser pouco ou muito; não há uma definição exata.

Contudo, é válido destacar que uma padronização das variáveis pode ter um viés negativo dentro da Sociolinguística Variacionista. Cada vez mais, as pesquisas têm se voltado para comunidades interioranas ou situações mais específicas de agrupamentos sociais — como as comunidades de prática (Eckert, 2022 [2012]) —, o que exige uma segmentação específica do público devido às particularidades de cada local. Nesse sentido, considerando que essas comunidades apresentam comportamentos sociais distintos, um trabalho mais individualizado é necessário, e a meta-análise não seria recomendada em tais situações. Um exemplo disso é a problematização recente acerca da variável “sexo/gênero”, que é costumeiramente binária, mas parece não representar adequadamente os estratos sociais (Freitag, 2015).

Nesse sentido, a prática de meta-análise pode ser relevante em pesquisas nos grandes centros urbanos. Nessas regiões, é mais fácil estruturar as variáveis como escolaridade e faixa etária da mesma maneira e conseguir encontrar informantes que se encaixem em todas as células. Esse trabalho padronizado favoreceria a meta-análise e permitiria desvelar os comportamentos linguísticos do português brasileiro em regiões urbanas.

Outro ponto a se destacar é que metade dos trabalhos que se encaixavam no tema precisaram ser descartados por não disponibilizarem os dados para as variáveis não significativas. Quase como uma tradição nos estudos sociolinguísticos, os autores reportam apenas os dados referentes às variáveis que o GoldVarb X não descarta. Consequentemente, a análise comparativa dos resultados das variáveis estatisticamente insignificantes é inviabilizada.

Assim, alguns princípios poderiam ser seguidos para facilitar a realização de mais meta-análises no futuro:

- 1) os trabalhos poderiam conter as palavras-chave da pesquisa no título para que sejam identificados na busca em bases de dados;
- 2) as variáveis independentes das pesquisas em centros urbanos poderiam ser construídas seguindo um padrão<sup>26</sup>; e
- 3) os trabalhos poderiam disponibilizar os dados das variáveis sem significância estatística (uma tabela com os dados numéricos e os valores em porcentagem já seria o suficiente).

---

<sup>26</sup>Embora se pudesse indagar acerca de quem definiria o padrão, esta resposta poderia ser dada nos próprios espaços de discussão entre os pesquisadores de diferentes estados. Em outras palavras, esse padrão seria uma convenção.

## 7 CONSIDERAÇÕES FINAIS

Esta monografia 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, tanto pelo escopo do trabalho ser menor quanto 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.

De início, foi feita uma revisão sistemática e foram selecionados dez trabalhos para análise geral da negação verbal e cinco estudos específicos que disponibilizavam dados para os três condicionamentos sociais. Depois disso, realizaram-se testes no RStudio com os dados disponibilizados e obtiveram-se resultados relevantes.

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.

## REFERÊNCIAS

- ALENCAR, Dulcelita Pereira Ribeiro de. **O Uso da Negação no Português Brasileiro: gramática e gramaticalização**. 2009. Dissertação (Mestrado em Língua Portuguesa) – Pontifícia Universidade Católica de São Paulo, São Paulo, 2009.
- ARAÚJO, Rerisson Cavalcante. **Negação Anafórica no Português Brasileiro: negação sentencial, negação enfática e negação de constituinte**. 2012. Tese (Doutorado em Linguística) – Universidade de São Paulo, São Paulo, 2012.
- BECHARA, Evanildo. **Moderna Gramática Portuguesa**. 39. ed. Rio de Janeiro: Nova Fronteira, 2019.
- BELLETTI, Thais de Freitas Mondini. Estratégias de Negação do Brasileiro na Crônica Nacional: uma contribuição para o ensino de português para estrangeiros. **Caleidoscópio**, v. 8, n. especial 2, 2020.
- BRAGA, Luciana; SILVA, Josenildo Ferreira Teófilo da. Análise Funcionalista das Estratégias de Negação do Português Oral Culto de Fortaleza: um estudo de caso. **Entrepalavras**, Fortaleza, v. 1, n. 1, ago./dez. 2011.
- BORTONI-RICARDO, Stella Maris. **Educação em língua materna: a sociolinguística na sala de aula**. São Paulo: Parábola, 2004.
- CASTRO, Savanna Souza de; VIEIRA, Julinara Silva; SOUSA, Valéria Viana; SILVA, Jorge Augusto Alves. Uma Análise Sociofuncionalista da Dupla Negação no Sertão da Ressaca. **Seminário de Pesquisa e Estudos Linguísticos**, v. 9, n. 1, 2017.
- CASTILHO, Ataliba de. **Nova gramática do português brasileiro**. São Paulo: Contexto, 2010.
- CAVALCANTE, Rerisson. **A Negação Pós-verbal no Português Brasileiro: análise descritiva e teórica de dialetos rurais de afro-descendentes**. 2007. Dissertação (Mestrado em Letras) – Universidade Federal da Bahia, Salvador, 2007.
- CAVALCANTE, Rerisson. Redobro da Negação Pós-verbal e Deslocamento à Direita. In: **Caderno de Squibs: Temas em estudos formais da linguagem**, v. 4, n. 1, 2018.
- 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, 2013.
- COELHO, Izete Lehmkuhl *et al.* **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.
- ECKERT, Penelope. As Três Ondas do Estudo da Variação: a emergência do significado no estudo da variação sociolinguística. **Organon**, Porto Alegre, v. 37, n. 73, p. 268-291, jan./jun. 2022 [2012]. DOI: 10.22456/2238-8915.122962.

FARACO, Carlos Alberto. Norma-padrão brasileira: desembaraçando alguns nós. *In*: BAGNO, Marcos (org.). **Linguística da Norma**. São Paulo: Edições Loyola, 2002.

FERREIRA JUNIOR, Moacir Natercio. **A Sintaxe da Negação em Configuração Imperativa no Português Brasileiro**. 2011. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 2011.

FREITAG, Raquel M. K. Idade: uma variável sociolinguística complexa. **Línguas & Letras**, v. 6, n. 11, p. 105-121, 2005.

FREITAG, Raquel M. K. (Re)Discutindo Sexo/Gênero na Sociolinguística. *In*: FREITAG, M. K.; SEVERO, C. G. (org.). **Mulheres, Linguagem e Poder - Estudos de Gênero na Sociolinguística Brasileira**. São Paulo: Blucher, 2015.

FREITAG, Raquel M. K. **Como fazer meta-análise com dados sociolinguísticos?**. [2021]. Disponível em: <https://rkofreitag.github.io/meta.html/>. Acesso em: 9 abr. 2023.

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.

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, 2013. DOI: <https://doi.org/10.5210%2Fojphi.v5i2.4623>.

GODOY, Mahayana C. Introdução aos modelos lineares mistos para os estudos da linguagem. **PsyArXiv**, 2019. DOI: <https://doi.org/10.17605/OSF.IO/9T8UR>.

GRAÇA, Beatriz; AVELAR, Maíra. Uma Análise Cognitivo-gestual de Gestos de Negação em Ocorrências Multimodais do Português. *In*: ENANPOLL, 35., 2020. [S. l.]. **Anais [...]**. [S. l.]: ANPOLL, 2020.

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

GONÇALVES, Rafael Teixeira; TEMPONI, Cristiane Namiuti; PACHECO, Vera. Um Estudo Sobre a Negação Sentencial no Português Brasileiro. **Seminário de Pesquisa em Estudos Linguísticos**, v. 8, n. 1, 2014.

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

LAMBERTI NUNES, Luana. A Idealização de um Teste de Aceitabilidade de Enunciados com Dupla Negação. *In*: SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS, 25., 2013. Porto Alegre. **Anais [...]**. Porto Alegre, 2013.

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



LIMA, Carlos Henrique Rocha. **Gramática Normativa da Língua Portuguesa**. 49. ed. Rio de Janeiro: José Olympio, 2011.

LIMA, Luana Santos de. **Restrições Pragmáticas às Estratégias de Negação Sentencial no Português Falado no Brasil**. Poster, [20--?].

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, 2020.

LOVATTO, P. A. *et al.* 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.

MARCELINO, Nara Juscely Minervino de Carvalho. **Sentenças de Negação com *É Ruim e Nem a Pau* no Português Brasileiro**. 2018. Tese (Doutorado em Linguística) – Universidade Federal do Rio Grande do Norte, Natal, 2018.

MARINHO, Adriano Souza. O Uso do Não no Português Brasileiro: a atenuação da negação como estratégia de polidez. **Revista Eletrônica do GEPPELE**, Ceará, ed. 7, v. 1, jan. 2019.

MEIRELES, Selma M. A Negação Sintática em Diálogos do Alemão e do Português do Brasil. **Pandaemonium germanicum**, n. 5, 2001.

MENDES, Eliana Amarante de Mendonça. Aspectos da Recusa (negação) na Conversação em Português-brasileiro. **Revista de Estudos Linguísticos**, Belo Horizonte, v. 1, n. 4, jan./jun. 1996.

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.

MORENO, António. **Aspectos da Negação no Português: uma abordagem enunciativa**. 2005. Tese (Doutorado em Linguística) – Universidade Nova de Lisboa, Lisboa, 2005.

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.

NAMIUTI, Cristiane. **Aspectos da História Gramatical do Português: interpolação, negação e mudança**. 2007. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas, Campinas, 2007.

NAMIUTI, Cristiane. Negação Sentencial na Diacronia do Português: variação com estabilidade. **Revista de Estudos Linguísticos**, Belo Horizonte, v. 16, n. 2, jul./dez. 2008.

NAMIUTI-TEMPONI, Cristiane. Para uma Visão Diacrônica da Negação em Português: derivação sintática e tipologia. **Sínteses**, 2009.

NAMIUTI, Cristiane; MIOTO, Carlos. Clíticos e Negação em Português: elementos para uma descrição gramatical. **Filologia e Linguística Portuguesa**, v. 16, n. especial, dez. 2014.

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

NHABINDE JÚNIOR, Gaspar. **Estudo Contrastivo sobre Aspectos da Negação em Português e Inglês**. 2011. Dissertação (Mestrado em Linguística) – Universidade Eduardo Mondlane, Maputo, 2011.

NOVAIS, V.; SIQUEIRA, M. A Variável Sexo/gênero no Português Falado no Sertão Alagoano. **Leitura**, Maceió, n. 66, p. 35-50, set./dez. 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, 2014.

OLIVEIRA, Raquel de Fátima Cruz; CARVALHO E SILVA, Thiago do Amaral de; PRATA, Patrícia. Aspectos da Negação do Português Arcaico dos Séculos XIII e XIV. **Língua, Literatura e Ensino**, v. 2, maio 2007.

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

PAIVA, Maria da Conceição. A Variável Gênero/Sexo. In: MOLLICA, C. M.; BRAGA, M. L. **Introdução à Sociolinguística**: o tratamento da variação. 4. ed. São Paulo: Contexto, 2020.

PASTOR-RAMON, Elena *et al.* Improving the Reliability of Literature Reviews: Detection of Retracted Articles through Academic Search Engines. **Eur. J. Investig. Health Psychol. Educ.**, v. 12, n. 5, 2022. DOI: <https://doi.org/10.3390/ejihpe12050034>.

PATIL, Indrajeet. Visualizations with statistical details: The 'ggstatsplot' approach. **Journal of Open Source Software**, v. 6, n. 61, p. 3167, 2021. DOI: 10.21105/joss.03167.

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

PESSOA, Regina Maria. Manifestações da Negação em Português: a negação implícita. **Alfa**, São Paulo, v. 29, 1985.

PINTO, Clara. Para a História da Negação: o minimizador *homem* no português antigo. **Estudos de Linguística Galega**, v. 7, 2015.

PINTO, Clara. **Minimizadores e Negação**: do português antigo ao português contemporâneo. Poster, 2017.

R CORE Team. R: A language and environment for statistical computing. Áustria: R Foundation for Statistical Computing, 2021.

REIMANN, Cristiana Aparecida; YACOVENCO, Lilian Coutinho. A Dupla Negação no Português Falado em Vitória/ES: traço da identidade linguística capixaba?. *In*: CONEL, 1., 2011. Vitória. **Anais [...]**. Vitória: UFES, 2011.

REIMANN, Cristiana Aparecida; YACOVENCO, Lilian Coutinho. A Negação no Português Falado em Vitória/ES: atuação de fatores discursivos e sintáticos. *In*: CONEL, 2., 2014. Vitória. **Anais [...]**. Vitória: UFES, 2014.

ROCHA, Rafael Stoppa. Negação Verbal no Português Paulistano: envelope de variação. **Estudos Linguísticos**, São Paulo, v. 41, n. 2, maio/ago. 2012.

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

SANTOS, Beatriz Fabiane Graça. **A Construção de um Repertório de Gestos de Negação para o Português Brasileiro**: uma proposta cognitivo-gestual. 2021. Dissertação (Mestrado em Linguística) – Universidade Estadual do Sudoeste da Bahia, Vitória da Conquista, 2021.

SCHERRE, Maria Marta Pereira. Sobre a Influência de Variáveis Sociais na Concordância Nominal. *In*: OLIVEIRA E SILVA, G. M.; SCHERRE, M. M. P. **Padrões Sociolinguísticos**: análise de fenômenos variáveis no português falado na cidade do Rio de Janeiro. Rio de Janeiro: Tempo Brasileiro, 1996.

SCHERRE, Maria Marta Pereira; YACOVENCO, Lilian Coutinho. A Variação Linguística e o Papel dos Fatores Sociais: o gênero do falante em foco. **Revista da ABRALIN**, v. eletrônico, n. especial, p. 121-146, 2011.

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) – Universidade Federal do Maranhão, São Luís, 2018.

SILVA, Edilson dos Santos. **As Estratégias de Negação Variantes no Inglês Falado**: um estudo comparativo com o português. 2008. Dissertação (Mestrado em Estudos da Linguagem) – Universidade Federal do Rio Grande do Norte, Natal, 2008.

SILVA, Luis Filipe. MELLO, Heliana. Investigação sobre a Negação Verbal no Português Brasileiro Através de Corpus de Fala Espontânea: restrições prosódico-informacionais. *In*: CONGRESSO BRASILEIRO DE PROSÓDIA, 5., 2015, Brasília. **Anais [...]**. Brasília, 2015.

SILVA, Luis Filipe Lima e. **Negação Verbal no Português Brasileiro**: aspectos teórico-metodológicos em estudo baseado em corpus. 2016. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Belo Horizonte, 2016.

SILVA, Vera Lúcia Paredes da. Relevância das variáveis linguísticas. *In*: MOLLICA, Maria Cecília; BRAGA, Maria Luiza. **Introdução à Sociolinguística**: o tratamento da variação. 4. ed. São Paulo: Contexto, 2020.

SILVEIRA, Regina Célia Pagliuchi da. Conteúdos Lingüísticos e Implícitos Culturais em Expressões da Negação, no Uso Efetivo do Português Brasileiro. *In*: JORNADA GELNE, 10., 2004, João Pessoa. **Anais [...]**. João Pessoa: GELNE, 2004.

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

SOUSA, Lílían Teixeira de. **Formas Reduzidas de Itens Negativos no Português Brasileiro**. 2007. Dissertação (Mestrado em Linguística) – Universidade Federal de Minas Gerais, Belo Horizonte, 2007.

SOUSA, Lílían Teixeira de. A Negação Sentencial e o Efeito de Bloqueio no Português Brasileiro. **Revista Recorte**, v. 9, n. 2, 2012.

SOUSA, Lílían Teixeira de. Sobre as Origens da Distinção entre Negação de Proposição e Negação de Evento no Português Brasileiro. **Estudos de Linguística Galega**, v. especial 1, 2018a.

SOUSA, Lílían Teixeira de. Sobre Negação e Tempo: um estudo de caso sobre o português brasileiro. **Caderno de Estudos Linguísticos**, Campinas, v. 60, n. 1, jan./abr. 2018b.

SOUZA, Paula da Costa. **A Dupla Negação Pré-verbal no Catalão e no Português Brasileiro**: história, variação e uso. 2017. Tese (Doutorado em Filologia e Língua Portuguesa) – Universidade de São Paulo, São Paulo, 2017.

VIECHTBAUER, W. Conducting meta-analyses in R with the metafor package. **Journal of Statistical Software**, v. 36, n. 3, p. 1-48, 2010. DOI: <https://doi.org/10.18637/jss.v036.i03>.

VIEIRA, N. M. T. **Monotongação de Ditongos Orais no Português Brasileiro**: uma revisão sistemática da literatura. 2021. Dissertação (Mestrado em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2021.

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

VITÓRIO, E. G. S. L. A. A Língua Usada no Sertão Alagoano: constituição da amostra. *In*: OLIVEIRA, A. A.; PAULA, A. S. (Orgs.). **Interfaces Sociolinguísticas**: análises variacionistas em Alagoas e Pernambuco. Arapiraca: Eduneal, 2020.

TOKARNIA, Mariana. Analfabetismo cai, mas Brasil ainda tem 11 milhões sem ler e escrever. **Agência Brasil**, 15 de julho de 2020. Disponível em: <https://agenciabrasil.ebc.com.br/educacao/noticia/2020-07/taxa-cai-levemente-mas-brasil-ainda-tem-11-milhoes-de-analfabetos>. Acesso em: 2 fev. 2023.

VENABLES, W. N.; RIPLEY, B. D. **Modern Applied Statistics with S**. 4. ed. New York: Springer, 2002.

WICKHAM, Hadley *et al.* Welcome to the tidyverse. **Journal of Open Source Software**, v. 4, n. 43, p. 1686, 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, 2016.

## APÊNDICE A – PACOTES UTILIZADOS E ANÁLISE GERAL DOS DADOS

```
# Libraries
library (tidyverse)
library (sjPlot)
library (nnet)
library (ggstatsplot)
#-----
#-----
#-----
#ANÁLISE GERAL
VD = c ("Negação Pré-verbal", "Dupla Negação", "Negação Pós-verbal")
pesquisas = c ("Lopes da Silva, 2020", "Nunes, 2014", "Nascimento, 2014", "Rocha, 2013",
               "Cavalcante, 2007", "Serra, 2018", "Reimann e Yacovenco, 2011",
               "Rocha, 2012", "Reimann e Yacovenco, 2014")
dadostodos = data.matrix (rbind (
  c (1819, 278, 7),
  c (616, 214, 13),
  c (1751, 478, 34),
  c (5279, 324, 4),
  c (1343, 568, 115),
  c (1009, 133, 17),
  c (721, 216, 42),
  c (940, 117, 4),
  c (1754, 478, 34)
))

rownames (dadostodos) = pesquisas
colnames (dadostodos) = VD
View (dadostodos)

dadostodos = as.data.frame.table (dadostodos)
dadostodos = countsToCases(as.data.frame (dadostodos))
names (dadostodos) = c ('pesquisas', 'VD')
dadostodos$pesquisas = factor (dadostodos$pesquisas, levels = c ("Lopes da Silva, 2020",
"Nunes, 2014", "Nascimento, 2014", "Rocha, 2013",
"Cavalcante, 2007", "Serra, 2018", "Reimann e
Yacovenco, 2011",
"Rocha, 2012", "Reimann e Yacovenco, 2014"))

head(dadostodos)
ggbarstats (x = VD, y =pesquisas, data=dadostodos,
            bf.message = F,
            xlab = "Pesquisas",
            legend.title = "Negação")
```

## APÊNDICE B – ANÁLISE DA VARIÁVEL SEXO/GÊNERO

### #ANÁLISE DA VARIÁVEL GÊNERO NO USO DA NEGAÇÃO DUPLA COM VALOR DE REFERÊNCIA "NEGAÇÃO PRÉ-VERBAL"

```
#Cavalcante (2007)
```

```
VD = c ("Negação Pré-verbal", "Dupla Negação")
```

```
genero = c ("masculino", 'feminino')
```

```
cav2007 = data.matrix (rbind (
```

```
  c (585, 205),
```

```
  c (759, 363)
```

```
))
```

```
rownames (cav2007) = genero
```

```
colnames (cav2007) = VD
```

```
View (cav2007)
```

```
cav2007.t = t(cav2007)
```

```
barplot (cav2007.t,
```

```
  beside = T,
```

```
  legend = T,
```

```
  cex.axis = 0.8,
```

```
  col = c ('aquamarine2', 'aliceblue'),
```

```
  cex.names = 0.8,
```

```
  ylim= c (0,1000),
```

```
  args.legend = list (x='topleft',
```

```
                      cex = 0.8, bty = 'n'),
```

```
  ylab = "Frequência de Uso")
```

```
cav2007 = as.data.frame.table (cav2007)
```

```
cav2007 = countsToCases(as.data.frame (cav2007))
```

```
names (cav2007) = c ('genero', 'VD')
```

```
cav2007$genero = factor (cav2007$genero, levels = c ('masculino', 'feminino'))
```

```
head(cav2007)
```

```
ggbarstats (x = VD, y =genero, data=cav2007,
```

```
  bf.message = F,
```

```
  xlab = "Gênero",
```

```
  legend.title = "Negação")
```

```
table (cav2007)
```

```
cav2007$VD = relevel (cav2007$VD, ref='Negação Pré-verbal')
```

```
modelo = glm (VD ~ genero, data=cav2007, family = 'binomial')
```

```
summary (modelo)
```

```
tab_model (modelo, transform = NULL)
```

```
#-----
```

```
#-----
```

```
#-----
```

```
# #Lopes da Silva (2020)
```

```
VD = c ("Negação Pré-verbal", "Dupla Negação")
```

```
genero = c ("masculino", 'feminino')
```

```

lopes2020 = data.matrix (rbind (
  c (817, 124),
  c (1002, 154)
))
rownames (lopes2020) = genero
colnames (lopes2020) = VD
View (lopes2020)

lopes2020.t = t(lopes2020)
barplot (lopes2020.t,
  beside = T,
  legend = T,
  cex.axis = 0.8,
  col = c ('aquamarine2', 'aliceblue'),
  cex.names = 0.8,
  ylim= c (0,1000),
  args.legend = list (x='topleft',
    cex = 0.8, bty = 'n'),
  ylab = "Frequência de Uso")

lopes2020 = as.data.frame.table (lopes2020)
lopes2020 = countsToCases(as.data.frame (lopes2020))
names (lopes2020) = c ('genero', 'VD')
lopes2020$genero = factor (lopes2020$genero, levels = c ('masculino', 'feminino'))
head(lopes2020)
ggbarstats (x = VD, y = genero, data=lopes2020,
  bf.message = F,
  xlab = "Gênero",
  legend.title = "Negação")
table (lopes2020)

lopes2020$VD = relevel (lopes2020$VD, ref='Negação Pré-verbal')
modelo2 = glm (VD ~ genero, data=lopes2020, family = 'binomial')
summary (modelo2)
tab_model (modelo2, transform = NULL)
#-----
#-----
#-----
# Nunes (2014)
VD = c ("Negação Pré-verbal", "Dupla Negação")
genero = c ("masculino", 'feminino')
nun2014 = data.matrix (rbind (
  c (264, 109),
  c (352, 105)
))
rownames (nun2014) = genero
colnames (nun2014) = VD
View (nun2014)

nun2014.t = t(nun2014)

```



```

barplot (nun2014.t,
        beside = T,
        legend = T,
        cex.axis = 0.8,
        col = c ('aquamarine2', 'aliceblue'),
        cex.names = 0.8,
        ylim= c (0,1000),
        args.legend = list (x='topleft',
                             cex = 0.8, bty = 'n'),
        ylab = "Frequência de Uso")

nun2014 = as.data.frame.table (nun2014)
nun2014 = countsToCases(as.data.frame (nun2014))
names (nun2014) = c ('genero', 'VD')
nun2014$genero = factor (nun2014$genero, levels = c ('masculino', 'feminino'))
head(nun2014)
ggbarstats (x = VD, y =genero, data=nun2014,
            bf.message = F,
            xlab = "Gênero",
            legend.title = "Negação")
table (nun2014)

nun2014$VD = relevel (nun2014$VD, ref='Negação Pré-verbal')
modelo3 = glm (VD ~ genero, data=nun2014, family = 'binomial')
summary (modelo3)
tab_model (modelo3, transform = NULL)
#-----
#-----
#-----
#Nascimento (2014)
VD = c ("Negação Pré-verbal", "Dupla Negação")
genero = c ("masculino", 'feminino')
nasc2014 = data.matrix (rbind (
  c (881, 238),
  c (870, 240)
))
rownames (nasc2014) = genero
colnames (nasc2014) = VD
View (nasc2014)

nasc2014.t = t(nasc2014)
barplot (nasc2014.t,
        beside = T,
        legend = T,
        cex.axis = 0.8,
        col = c ('aquamarine2', 'aliceblue'),
        cex.names = 0.8,
        ylim= c (0,1000),
        args.legend = list (x='topleft',
                             cex = 0.8, bty = 'n'),

```

```

ylab = "Frequência de Uso")

nasc2014 = as.data.frame.table (nasc2014)
nasc2014 = countsToCases(as.data.frame (nasc2014))
names (nasc2014) = c ('genero', 'VD')
nasc2014$genero = factor (nasc2014$genero, levels = c ('masculino', 'feminino'))
head(nasc2014)
ggbarstats (x = VD, y =genero, data=nasc2014,
             bf.message = F,
             xlab = "Gênero",
             legend.title = "Negação")
table (nasc2014)

nasc2014$VD = relevel (nasc2014$VD, ref='Negação Pré-verbal')
modelo4 = glm (VD ~ genero, data=nasc2014, family='binomial')
summary (modelo4)
tab_model (modelo4, transform = NULL)
#-----
#-----
#-----
# Rocha (2013)
VD = c ("Negação Pré-verbal", "Dupla Negação")
genero = c ("masculino", 'feminino')
roc2013 = data.matrix (rbind (
  c (2434, 146),
  c (2845, 178)
))
rownames (roc2013) = genero
colnames (roc2013) = VD
View (roc2013)

roc2013.t = t(roc2013)
barplot (roc2013.t,
         beside = T,
         legend = T,
         cex.axis = 0.8,
         col = c ('aquamarine2', "aliceblue"),
         cex.names = 0.8,
         ylim= c (0,1000),
         args.legend = list (x='topleft',
                             cex = 0.8, bty = 'n'),
         ylab = "Frequência de Uso")

roc2013 = as.data.frame.table (roc2013)
roc2013 = countsToCases(as.data.frame (roc2013))
names (roc2013) = c ('genero', 'VD')
roc2013$genero = factor (roc2013$genero, levels = c ('masculino', 'feminino'))
head(roc2013)
ggbarstats (x = VD, y =genero, data=roc2013,
             bf.message = F,

```

```

      xlab = "Gênero",
      legend.title = "Negação")
table (roc2013)

roc2013$VD = relevel (roc2013$VD, ref='Negação Pré-verbal')
modelo5 = glm (VD ~ genero, data=roc2013, family = 'binomial')
summary (modelo5)
tab_model (modelo5, transform = NULL)
#-----
#-----
#-----
#METANÁLISE
df <-
  structure(
    list(
      term = structure(
        1:5,
        .Label = c("Sudeste \n (Nunes, 2014)", "Nordeste \n (Lopes da Silva, 2020)", "Nordeste
\n (Cavalcante, 2007)",
        "Sudeste \n (Nascimento, 2014)", "Sudeste \n (Rocha, 2013)"),
        class = "factor"
      ),
      estimate = c(
        -0.3251,
        0.01256,
        0.31100,
        0.02093,
        0.04215
      ),
      std.error = c(
        0.1591,
        0.12954,
        0.10325,
        0.10321,
        0.11502
      ),
      statistic = c(
        -2.043,
        0.097,
        3.012,
        0.203,
        0.366
      ),
      p.value = c(
        0.0411,
        0.923,
        0.00259,
        0.839,
        0.714
      ),

```

```

df.residual = c(
  828,
  2095,
  1910,
  2227,
  5601
)
),
row.names = c(NA, -5L),
class = c("tbl_df", "tbl", "data.frame")
)
### gerar a visualização gráfica:
ggcoefstats(
  x = df,
  statistic = "z",
  sort = "ascending",
  ylab = "Amostras",
  xlab = "Coeficiente de regressão",
  caption = "Meta-análise"
)

```

## #ANÁLISE DA VARIÁVEL GÊNERO NO USO DA NEGAÇÃO PÓS-VERBAL COM VALOR DE REFERÊNCIA "NEGAÇÃO DUPLA"

```

# Cavalcante (2007)
VD = c ("Dupla Negação", "Negação Pós-verbal")
genero = c ("masculino", 'feminino')
cav2007.2 = data.matrix (rbind (
  c (205, 61),
  c (363, 53)
))
rownames (cav2007.2) = genero
colnames (cav2007.2) = VD
View (cav2007.2)

cav2007.2.t = t(cav2007.2)
barplot (cav2007.2.t,
  beside = T,
  legend = T,
  cex.axis = 0.8,
  col = c ('aquamarine2', "aliceblue"),
  cex.names = 0.8,
  ylim= c (0,1000),
  args.legend = list (x='topleft',
    cex = 0.8, bty = 'n'),
  ylab = "Frequência de Uso")

```

```
cav2007.2 = as.data.frame.table (cav2007.2)
cav2007.2 = countsToCases(as.data.frame (cav2007.2))
names (cav2007.2) = c ('genero', 'VD')
cav2007.2$genero = factor (cav2007.2$genero, levels = c ('masculino', 'feminino'))
head(cav2007.2)
ggbarstats (x = VD, y =genero, data=cav2007.2,
             bf.message = F,
             xlab = "Gênero",
             legend.title = "Negação")
table (cav2007.2)
```

```
cav2007.2$VD = relevel (cav2007.2$VD, ref='Dupla Negação')
modelo.1 = glm (VD ~ genero, data=cav2007.2, family='binomial')
summary (modelo.1)
tab_model (modelo.1, transform = NULL)
#-----
#-----
#-----
# Nunes (2014)
VD = c ("Dupla Negação", "Negação Pós-verbal")
genero = c ("masculino", 'feminino')
nun2014.2 = data.matrix (rbind (
  c (109, 8),
  c (105, 5)
))
rownames (nun2014.2) = genero
colnames (nun2014.2) = VD
View (nun2014.2)
```

```
nun2014.2.t = t(nun2014.2)
barplot (nun2014.2.t,
         beside = T,
         legend = T,
         cex.axis = 0.8,
         col = c ('aquamarine2', "aliceblue"),
         cex.names = 0.8,
         ylim= c (0,1000),
         args.legend = list (x='topleft',
                             cex = 0.8, bty = 'n'),
         ylab = "Frequência de Uso")
```

```
nun2014.2 = as.data.frame.table (nun2014.2)
nun2014.2 = countsToCases(as.data.frame (nun2014.2))
names (nun2014.2) = c ('genero', 'VD')
nun2014.2$genero = factor (nun2014.2$genero, levels = c ('masculino', 'feminino'))
head(nun2014.2)
ggbarstats (x = VD, y =genero, data=nun2014.2,
             bf.message = F,
             xlab = "Gênero",
             legend.title = "Negação")
```

```

table (nun2014.2)

nun2014.2$VD = relevel (nun2014.2$VD, ref='Dupla Negação')
modelo.2 = glm (VD ~ genero, data=nun2014.2, family='binomial')
summary (modelo.2)
tab_model (modelo.2, transform = NULL)
#-----
#-----
#-----
# Nascimento (2014)
VD = c ("Dupla Negação", "Negação Pós-verbal")
genero = c ("masculino", 'feminino')
nasc2014.2 = data.matrix (rbind (
  c (238, 18),
  c (240, 16)
))
rownames (nasc2014.2) = genero
colnames (nasc2014.2) = VD
View (nasc2014.2)

nasc2014.2.t = t(nasc2014.2)
barplot (nasc2014.2.t,
  beside = T,
  legend = T,
  cex.axis = 0.8,
  col = c ('aquamarine2', 'aliceblue'),
  cex.names = 0.8,
  ylim= c (0,1000),
  args.legend = list (x='topleft',
    cex = 0.8, bty = 'n'),
  ylab = "Frequência de Uso")

nasc2014.2 = as.data.frame.table (nasc2014.2)
nasc2014.2 = countsToCases(as.data.frame (nasc2014.2))
names (nasc2014.2) = c ('genero', 'VD')
nasc2014.2$genero = factor (nasc2014.2$genero, levels = c ('masculino', 'feminino'))
head(nasc2014.2)
ggbarstats (x = VD, y =genero, data=nasc2014.2,
  bf.message = F,
  xlab = "Gênero",
  legend.title = "Negação")
table (nasc2014.2)

nasc2014.2$VD = relevel (nasc2014.2$VD, ref='Dupla Negação')
modelo.3 = glm (VD ~ genero, data=nasc2014.2, family='binomial')
summary (modelo.3)
tab_model (modelo.3, transform = NULL)
#-----
#-----
#-----

```

```

#METANÁLISE
df2 <-
  structure(
    list(
      term = structure(
        1:3,
        .Label = c("Sudeste \n (Nunes, 2014)", "Nordeste \n (Cavalcante, 2007)", "Sudeste \n
(Nascimento, 2014)"),
        class = "factor"
      ),
      estimate = c(
        -0.4326,
        -0.7120,
        -0.1262
      ),
      std.error = c(
        0.5862,
        0.2071,
        0.3556
      ),
      statistic = c(
        -0.738,
        -3.438,
        -0.355
      ),
      p.value = c(
        0.461,
        0.000587,
        0.723
      ),
      df.residual = c(
        225,
        680,
        510
      )
    ),
    row.names = c(NA, -3L),
    class = c("tbl_df", "tbl", "data.frame")
  )
### gerar a visualização gráfica:
ggstatsplot::ggcoefstats(
  x = df2,
  statistic = "z",
  sort = "ascending",
  ylab = "Amostras",
  xlab = "Coeficiente de regressão",
  caption = "Meta-análise"
)

```

## APÊNDICE C – ANÁLISE DA VARIÁVEL ESCOLARIDADE

```
#ANÁLISE DA VARIÁVEL ESCOLARIDADE
```

```
# Cavalcante (2007)
```

```
VD = c ("Negação Pré-verbal", "Dupla Negação", "Negação Pós-verbal")
```

```
escolaridade = c ("Semi-Alfabetizados", 'Analfabetos')
```

```
cav2007esc = data.matrix (rbind (
```

```
  c (884, 302, 64),
```

```
  c (460, 266, 50)
```

```
))
```

```
rownames (cav2007esc) = escolaridade
```

```
colnames (cav2007esc) = VD
```

```
View (cav2007esc)
```

```
cav2007esc.t = t(cav2007esc)
```

```
barplot (cav2007esc.t,
```

```
  beside = T,
```

```
  legend = T,
```

```
  cex.axis = 0.8,
```

```
  col = c ('aquamarine2', 'aliceblue', 'azure4'),
```

```
  cex.names = 0.8,
```

```
  ylim= c (0,1000),
```

```
  args.legend = list (x='topleft',
```

```
                      cex = 0.8, bty = 'n'),
```

```
  ylab = "Frequência de Uso")
```

```
cav2007esc = as.data.frame.table (cav2007esc)
```

```
cav2007esc = countsToCases(as.data.frame (cav2007esc))
```

```
names (cav2007esc) = c ('escolaridade', 'VD')
```

```
cav2007esc$escolaridade = factor (cav2007esc$escolaridade, levels = c ('Semi-Alfabetizados',  
'Analfabetos'))
```

```
head(cav2007esc)
```

```
tail (cav2007esc)
```

```
ggbarstats (x = VD, y =escolaridade, data=cav2007esc,
```

```
  bf.message = F,
```

```
  xlab = "Escolaridade",
```

```
  legend.title = "Negação")
```

```
table (cav2007esc)
```

```
cav2007esc$VD = relevel (cav2007esc$VD, ref='Negação Pré-verbal')
```

```
mod.esc1 = multinom (VD ~ escolaridade, data=cav2007esc)
```

```
summary (mod.esc1)
```

```
tab_model (mod.esc1, transform = NULL)
```

```
#-----
```

```
#-----
```

```
#-----
```

```
#Lopes da Silva (2020)
```

```
VD = c ("Negação Pré-verbal", "Dupla Negação")
```

```
escolaridade = c ("Ensino Médio", 'Ensino Superior')
```



```

lopes2020esc = data.matrix (rbind (
  c (852, 136),
  c (967, 142)
))
rownames (lopes2020esc) = escolaridade
colnames (lopes2020esc) = VD
View (lopes2020esc)

lopes2020esc.t = t(lopes2020esc)
barplot (lopes2020esc.t,
  beside = T,
  legend = T,
  cex.axis = 0.8,
  col = c ('aquamarine2', 'aliceblue'),
  cex.names = 0.8,
  ylim= c (0,1000),
  args.legend = list (x='topleft',
    cex = 0.8, bty = 'n'),
  ylab = "Frequência de Uso")

lopes2020esc = as.data.frame.table (lopes2020esc)
lopes2020esc = countsToCases(as.data.frame (lopes2020esc))
names (lopes2020esc) = c ('escolaridade', 'VD')
lopes2020esc$escolaridade = factor (lopes2020esc$escolaridade, levels = c ('Ensino Médio',
'Ensino Superior'))
head(lopes2020esc)
ggbarstats (x = VD, y =escolaridade, data=lopes2020esc,
  bf.message = F,
  xlab = "Escolaridade",
  legend.title = "Negação")
table (lopes2020esc)

lopes2020esc$VD = relevel (lopes2020esc$VD, ref='Negação Pré-verbal')
mod.esc2 = glm (VD ~ escolaridade, data=lopes2020esc, family = 'binomial')
summary (mod.esc2)
tab_model (mod.esc2, transform = NULL)
#-----
#-----
#-----
# Nunes (2014)
VD = c ("Negação Pré-verbal", "Dupla Negação", "Negação Pós-verbal")
escolaridade = c ("5-8 anos", '9-11 anos')
nun2014esc = data.matrix (rbind (
  c (324, 131, 10),
  c (292, 83, 3)
))
rownames (nun2014esc) = escolaridade
colnames (nun2014esc) = VD
View (nun2014esc)

```

```

nun2014esc.t = t(nun2014esc)
barplot (nun2014esc.t,
        beside = T,
        legend = T,
        cex.axis = 0.8,
        col = c ('aquamarine2', 'aliceblue', 'azure4'),
        cex.names = 0.8,
        ylim= c (0,1000),
        args.legend = list (x='topleft',
                           cex = 0.8, bty = 'n'),
        ylab = "Frequência de Uso")

nun2014esc = as.data.frame.table (nun2014esc)
nun2014esc = countsToCases(as.data.frame (nun2014esc))
names (nun2014esc) = c ('escolaridade', 'VD')
nun2014esc$escolaridade = factor (nun2014esc$escolaridade, levels = c ('5-8 anos', '9-11
anos'))
head(nun2014esc)
ggbarstats (x = VD, y =escolaridade, data=nun2014esc,
            bf.message = F,
            xlab = "Gênero",
            legend.title = "Negação")
table (nun2014esc)

nun2014esc$VD = relevel (nun2014esc$VD, ref='Negação Pré-verbal')
mod.esc3 = multinom (VD ~ escolaridade, data=nun2014esc)
summary (mod.esc3)
tab_model (mod.esc3, transform = NULL)
#-----
#-----
#-----
# Nascimento (2014)
VD = c ("Negação Pré-verbal", "Dupla Negação", "Negação Pós-verbal")
escolaridade = c ("Ensino Fundamental", 'Ensino Médio', "Ensino Universitário")
nasc2014esc = data.matrix (rbind (
  c (544, 178, 16),
  c (695, 166, 6),
  c (512, 134, 12)
))
rownames (nasc2014esc) = escolaridade
colnames (nasc2014esc) = VD
View (nasc2014esc)

nasc2014esc.t = t(nasc2014esc)
barplot (nasc2014esc.t,
        beside = T,
        legend = T,
        cex.axis = 0.8,
        col = c ('aquamarine2', 'aliceblue', 'azure4'),
        cex.names = 0.8,

```

```

ylim= c (0,1000),
args.legend = list (x='topleft',
                    cex = 0.8, bty = 'n'),
ylab = "Frequência de Uso")

nasc2014esc = as.data.frame.table (nasc2014esc)
nasc2014esc = countsToCases(as.data.frame (nasc2014esc))
names (nasc2014esc) = c ('escolaridade', 'VD')
nasc2014esc$escolaridade = factor (nasc2014esc$escolaridade, levels = c ('Ensino
Fundamental', 'Ensino Médio', "Ensino Universitário"))
head(nasc2014esc)
ggbarstats (x = VD, y =escolaridade, data=nasc2014esc,
            bf.message = F,
            xlab = "Escolaridade",
            legend.title = "Negação")
table (nasc2014esc)

nasc2014esc$VD = relevel (nasc2014esc$VD, ref='Negação Pré-verbal')
mod.esc4 = multinom (VD ~ escolaridade, data=nasc2014esc)
summary (mod.esc4)
tab_model (mod.esc4, transform = NULL)
#-----
#-----
#-----
# Rocha (2013)
VD = c ("Negação Pré-verbal", "Dupla Negação")
escolaridade = c ("Ensino Médio", 'Ensino Superior')
roc2013esc = data.matrix (rbind (
  c (2491, 202),
  c (2788, 122)
))
rownames (roc2013esc) = escolaridade
colnames (roc2013esc) = VD
View (roc2013esc)

roc2013esc.t = t(roc2013esc)
barplot (roc2013esc.t,
        beside = T,
        legend = T,
        cex.axis = 0.8,
        col = c ('aquamarine2', "aliceblue"),
        cex.names = 0.8,
        ylim= c (0,3000),
        args.legend = list (x='topleft',
                            cex = 0.8, bty = 'n'),
        ylab = "Frequência de Uso")

roc2013esc = as.data.frame.table (roc2013esc)
roc2013esc = countsToCases(as.data.frame (roc2013esc))
names (roc2013esc) = c ('escolaridade', 'VD')

```

```
roc2013esc$escolaridade = factor (roc2013esc$escolaridade, levels = c ('Ensino Médio',  
'Ensino Superior'))  
head(roc2013esc)  
ggbarstats (x = VD, y =escolaridade, data=roc2013esc,  
            bf.message = F,  
            xlab = "Escolaridade",  
            legend.title = "Negação")  
table (roc2013esc)  
  
roc2013esc$VD = relevel (roc2013esc$VD, ref='Negação Pré-verbal')  
mod.esc5 = glm (VD ~ escolaridade, data=roc2013esc, family = 'binomial')  
summary (mod.esc5)  
tab_model (mod.esc5, transform = NULL)
```

## APÊNDICE D – ANÁLISE DA VARIÁVEL FAIXA ETÁRIA

### #ANÁLISE DA VARIÁVEL FAIXA ETÁRIA

# Cavalcante (2007)

VD = c ("Negação Pré-verbal", "Dupla Negação", "Negação Pós-verbal")

faixa.etaria = c ("20-40 anos", '41-60 anos', "+60 anos")

cav2007fe = data.matrix (rbind (

  c (411, 182, 45),

  c (556, 198, 22),

  c (377, 188, 47)

))

rownames (cav2007fe) = faixa.etaria

colnames (cav2007fe) = VD

View (cav2007fe)

cav2007fe.t = t(cav2007fe)

barplot (cav2007fe.t,

  beside = T,

  legend = T,

  cex.axis = 0.8,

  col = c ('aquamarine2', "aliceblue", "azure4"),

  cex.names = 0.8,

  ylim= c (0,1000),

  args.legend = list (x='topleft',

                      cex = 0.8, bty = 'n'),

  ylab = "Frequência de Uso")

cav2007fe = as.data.frame.table (cav2007fe)

cav2007fe = countsToCases(as.data.frame (cav2007fe))

names (cav2007fe) = c ('faixa.etaria', 'VD')

cav2007fe\$faixa.etaria = factor (cav2007fe\$faixa.etaria, levels = c ("20-40 anos", '41-60 anos', "+60 anos"))

head(cav2007fe)

ggbarstats (x = VD, y =faixa.etaria, data=cav2007fe,

  bf.message = F,

  xlab = "Faixa Etária",

  legend.title = "Negação")

table (cav2007fe)

cav2007fe\$VD = relevel (cav2007fe\$VD, ref='Negação Pré-verbal')

mod.fe1 = multinom (VD ~ faixa.etaria, data=cav2007fe)

summary (mod.fe1)

tab\_model (mod.fe1, transform = NULL)

#-----

#-----

#-----

#Lopes da Silva (2020)

VD = c ("Negação Pré-verbal", "Dupla Negação")

faixa.etaria = c ("-30 anos", '+30 anos')

```

lopes2020fe = data.matrix (rbind (
  c (953, 124),
  c (866, 154)
))
rownames (lopes2020fe) = faixa.etaria
colnames (lopes2020fe) = VD
View (lopes2020fe)

lopes2020fe.t = t(lopes2020fe)
barplot (lopes2020fe.t,
  beside = T,
  legend = T,
  cex.axis = 0.8,
  col = c ('aquamarine2', 'aliceblue'),
  cex.names = 0.8,
  ylim= c (0,1000),
  args.legend = list (x='topleft',
    cex = 0.8, bty = 'n'),
  ylab = "Frequência de Uso")

lopes2020fe = as.data.frame.table (lopes2020fe)
lopes2020fe = countsToCases(as.data.frame (lopes2020fe))
names (lopes2020fe) = c ('faixa.etaria', 'VD')
lopes2020fe$faixa.etaria = factor (lopes2020fe$faixa.etaria, levels = c ("-30 anos", '+30 anos'))
head(lopes2020fe)
ggbarstats (x = VD, y =faixa.etaria, data=lopes2020fe,
  bf.message = F,
  xlab = "Faixa Etária",
  legend.title = "Negação")
table (lopes2020fe)

lopes2020fe$VD = relevel (lopes2020fe$VD, ref='Negação Pré-verbal')
mod.fe2 = glm (VD ~ faixa.etaria, data=lopes2020fe, family = 'binomial')
summary (mod.fe2)
tab_model (mod.fe2, transform = NULL)
#-----
#-----
#-----
# Nunes (2014)
VD = c ("Negação Pré-verbal", "Dupla Negação", "Negação Pós-verbal")
faixa.etaria = c ("15-25 anos", '26-50 anos')
nun2014fe = data.matrix (rbind (
  c (235, 116, 3),
  c (381, 98, 10)
))
rownames (nun2014fe) = faixa.etaria
colnames (nun2014fe) = VD
View (nun2014fe)

nun2014fe.t = t(nun2014fe)

```

```

barplot (nun2014fe.t,
        beside = T,
        legend = T,
        cex.axis = 0.8,
        col = c ('aquamarine2', 'aliceblue', 'azure4'),
        cex.names = 0.8,
        ylim= c (0,1000),
        args.legend = list (x='topleft',
                             cex = 0.8, bty = 'n'),
        ylab = "Frequência de Uso")

nun2014fe = as.data.frame.table (nun2014fe)
nun2014fe = countsToCases(as.data.frame (nun2014fe))
names (nun2014fe) = c ('faixa.etaria', 'VD')
nun2014fe$faixa.etaria = factor (nun2014fe$faixa.etaria, levels = c ("15-25 anos", '26-50
anos'))
head(nun2014fe)
ggbarstats (x = VD, y =faixa.etaria, data=nun2014fe,
            bf.message = F,
            xlab = "Faixa Etária",
            legend.title = "Negação")
table (nun2014fe)

nun2014fe$VD = relevel (nun2014fe$VD, ref='Negação Pré-verbal')
mod.fe3 = multinom (VD ~ faixa.etaria, data=nun2014fe)
summary (mod.fe3)
tab_model (mod.fe3, transform = NULL)
#-----
#-----
#-----
# Nascimento (2014)
VD = c ("Negação Pré-verbal", "Dupla Negação", "Negação Pós-verbal")
faixa.etaria = c ("15-25 anos", '26-49 anos', '>49 anos')
nasc2014fe = data.matrix (rbind (
  c (566, 164, 15),
  c (557, 146, 7),
  c (628, 168, 12)
))
rownames (nasc2014fe) = faixa.etaria
colnames (nasc2014fe) = VD
View (nasc2014fe)

nasc2014fe.t = t(nasc2014fe)
barplot (nasc2014fe.t,
        beside = T,
        legend = T,
        cex.axis = 0.8,
        col = c ('aquamarine2', 'aliceblue', 'azure4'),
        cex.names = 0.8,
        ylim= c (0,1000),

```

```

args.legend = list (x='topleft',
                    cex = 0.8, bty = 'n'),
ylab = "Frequência de Uso")

nasc2014fe = as.data.frame.table (nasc2014fe)
nasc2014fe = countsToCases(as.data.frame (nasc2014fe))
names (nasc2014fe) = c ('faixa.etaria', 'VD')
nasc2014fe$faixa.etaria = factor (nasc2014fe$faixa.etaria, levels = c ("15-25 anos", '26-49
anos', '>49 anos'))
head(nasc2014fe)
ggbarstats (x = VD, y =faixa.etaria, data=nasc2014fe,
            bf.message = F,
            xlab = "Faixa Etária",
            legend.title = "Negação")
table (nasc2014fe)

nasc2014fe$VD = relevel (nasc2014fe$VD, ref='Negação Pré-verbal')
mod.fe4 = multinom (VD ~ faixa.etaria, data=nasc2014fe)
summary (mod.fe4)
tab_model (mod.fe4, transform = NULL)
#-----
#-----
#-----
# Rocha (2013)
VD = c ("Negação Pré-verbal", "Dupla Negação")
faixa.etaria = c ("18-35 anos", '36-55 anos', '>56 anos')
roc2013fe = data.matrix (rbind (
  c (1917, 101),
  c (1683, 106),
  c (1679, 117)
))
rownames (roc2013fe) = faixa.etaria
colnames (roc2013fe) = VD
View (roc2013fe)

roc2013fe.t = t(roc2013fe)
barplot (roc2013fe.t,
        beside = T,
        legend = T,
        cex.axis = 0.8,
        col = c ('aquamarine2', "aliceblue"),
        cex.names = 0.8,
        ylim= c (0,2500),
        args.legend = list (x='topleft',
                            cex = 0.8, bty = 'n'),
        ylab = "Frequência de Uso")

roc2013fe = as.data.frame.table (roc2013fe)
roc2013fe = countsToCases(as.data.frame (roc2013fe))
names (roc2013fe) = c ('faixa.etaria', 'VD')

```



```

roc2013fe$faixa.etaria = factor (roc2013fe$faixa.etaria, levels = c ("18-35 anos", '36-55 anos',
'>56 anos'))
head(roc2013fe)
ggbarstats (x = VD, y =faixa.etaria, data=roc2013fe,
             bf.message = F,
             xlab = "Faixa Etária",
             legend.title = "Negação")
table (roc2013fe)

roc2013fe$VD = relevel (roc2013fe$VD, ref='Negação Pré-verbal')
mod.fe5 = glm (VD ~ faixa.etaria, data=roc2013fe, family = 'binomial')
summary (mod.fe5)
tab_model (mod.fe5, transform = NULL)
#-----
#-----
#-----

```