

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

THOMAZ TORRES TEIXEIRA

**VARIAÇÃO LINGUÍSTICA ENTRE *NÓS* E *A GENTE* NA FUNÇÃO DE SUJEITO E
O GRADIENTE DE ESTILOS CONTEXTUAIS EM ENTREVISTAS
SOCIOLINGUÍSTICAS DE PORTO ALEGRE (RS)**

PORTO ALEGRE

2021

THOMAZ TORRES TEIXEIRA

**VARIAÇÃO LINGUÍSTICA ENTRE *NÓS* E *A GENTE* NA FUNÇÃO DE SUJEITO E
O GRADIENTE DE ESTILOS CONTEXTUAIS EM ENTREVISTAS
SOCIOLINGUÍSTICAS DE PORTO ALEGRE (RS)**

Monografia apresentada ao Instituto de Letras
da Universidade Federal do Rio Grande do Sul
como requisito parcial para a conclusão do
curso de Bacharelado em Letras.

Profa. Dra. Elisa Battisti

Orientadora

PORTO ALEGRE

2021

CIP - Catalogação na Publicação

Teixeira, Thomaz Torres
VARIÇÃO LINGUÍSTICA ENTRE NÓS E A GENTE NA FUNÇÃO
DE SUJEITO E O GRADIENTE DE ESTILOS CONTEXTUAIS EM
ENTREVISTAS SOCIOLINGUÍSTICAS DE PORTO ALEGRE (RS) /
Thomaz Torres Teixeira. -- 2021.
62 f.
Orientadora: Elisa Battisti.

Trabalho de conclusão de curso (Graduação) --
Universidade Federal do Rio Grande do Sul, Instituto
de Letras, Bacharelado em Letras: Tradutor Português e
Inglês, Porto Alegre, BR-RS, 2021.

1. Nós e a gente. 2. Sociolinguística. 3. Variação
Linguística. 4. Gramaticalização. 5. Estilo. I.
Battisti, Elisa, orient. II. Título.

AGRADECIMENTOS

Nestas linhas busco fazer um sincero agradecimento a todos aqueles que, de alguma forma, contribuíram para que eu pudesse chegar até aqui, pois uma realização como essa é uma obra feita por várias mãos.

Agradeço, em primeiro lugar, aos meus pais, Rudá e Suzana, por todos os ensinamentos, pelo esforço para proporcionar sempre o melhor possível para mim e meus irmãos e, principalmente, por nunca terem desistido de ir até o limite em busca das melhores condições para uma pessoa PCD, como eu. Muito obrigado. Agradeço, também, ao meu irmão, Marcelo, por ter sido uma inspiração durante toda minha vida, como pessoa, estudante e, além disso, pela amizade que construímos ao longo do tempo.

À professora Elisa Battisti, pela orientação neste trabalho, e por ter aberto as portas para mim em seu grupo de pesquisa, assim me proporcionando a possibilidade de conhecer um universo infinito. Agradeço por toda a troca de conhecimento, pela parceria e pela experiência maravilhosa de Iniciação Científica que proporcionou não só a mim, mas a todas as pessoas que se envolvem com o LínguaPOA.

Ao CNPq que, em meio a tantos ataques, segue resistindo pela sobrevivência da pesquisa no Brasil. À UFRGS, pelo ensino público, gratuito e de qualidade. Ainda, fico na torcida para que, ano após ano, mais e mais pessoas de grupos minoritários perambularem pelos campi desta universidade.

A todos(as) os professores(as) que fizeram parte da minha trajetória, desde a primeira série até o último semestre da graduação, o meu profundo agradecimento. Se fui intelectualmente apto a elaborar este trabalho, é pela contribuição de cada um de vocês, que sofrem todos os dias com o sucateamento da profissão neste país.

Ao meu trio de amigos do Bacharelado, Caroline, Júlia e Anderson, pelas risadas, pela parceria nos momentos bons e ruins e por terem proporcionado um ambiente muito mais leve durante todos esses anos. Agradeço, também, a todos que foram meus colegas de IC. Saibam que este trabalho tem muito da contribuição de vocês.

Um Trabalho de Conclusão de Curso, formalmente, tem apenas um autor, mas essa linda conquista é de todos nós. Muito obrigado por tudo!

RESUMO

Este trabalho investiga a variação entre as formas pronominais de primeira pessoa do plural *nós* e *a gente*, na função de sujeito, no português de Porto Alegre. Realiza-se análise em tempo real (LABOV, 1994) e controla-se a alternância estilística nas entrevistas sociolinguísticas de que são levantados os dados. Tem-se como base teórica a Teoria da Variação e Mudança (LABOV, 2008), além de fundamentos complementares como a Gramaticalização (GONÇALVES *et al.*, 2007), os Estilos Contextuais (LABOV, 2008) e a Árvore de Decisão (LABOV, 2001). Estudos anteriores acerca desta variável (LOPES, 1998; VIANNA e LOPES, 2012; BORGES, 2004; ZILLES, 2007; SOUZA e BOTASSINI, 2009; VITÓRIO, 2017) sugerem que a variação entre os pronomes *nós* e *a gente* se trata de uma mudança em progresso generalizada nas variedades do português brasileiro. Nesta investigação, são aplicados dois procedimentos atrelados à sociolinguística variacionista: a análise de regra variável, que implica levantar dados de fala de entrevistas sociolinguísticas, codificá-los e submetê-los à análise estatística de regressão logística; e os critérios da Árvore de Decisão, para identificar trechos de fala monitorada vs. fala espontânea nas entrevistas e codificar os dados para a variável *Estilo*. Tais procedimentos foram realizados em dados das amostras do VARSUL (1990) e LínguaPOA (2015-2019). A análise estatística foi efetuada no software estatístico R, interface RStudio, com as amostras de 12 informantes do VARSUL e 12 informantes do LínguaPOA. Os resultados para as três variáveis previsoras linguísticas – *Marca Morfêmica*, *Paralelismo Formal*, *Estilo* – e duas variáveis previsoras sociais – *Gênero*, *Faixa Etária* – controladas apontam para um aumento significativo na realização da variante *a gente* na amostra mais recente, em comparação com os resultados da amostra anterior (65% nos dados do VARSUL vs. 87% nos dados do LínguaPOA). A variação entre os pronomes *nós* e *a gente* é favorecida pela variável *Paralelismo Formal*, nas duas amostras, e pela variável *Gênero*, apenas na amostra do VARSUL. Em relação à alternância estilística, os resultados do cruzamento entre *Gênero*, *Estilo* e a variável-resposta *nós* apontaram, na amostra mais antiga (VARSUL), uma preferência dos homens pelo uso da variante *nós*, bem como um aumento da realização por parte de ambos os gêneros deste pronome nos contextos [+monitorados]. Já na amostra mais recente (LínguaPOA), tem-se a hipótese de que o alto percentual de realização de *a gente* (87%) e a ausência de informantes de ensino fundamental minimizem a influência da alternância estilística mesmo nos contextos de fala monitorada.

Palavras-chave: Nós e a gente. Sociolinguística. Variação Linguística. Gramaticalização. Estilo.

ABSTRACT

This work investigates the variation between the first-person plural pronouns *nós* and *a gente*, in the subject function, in Porto Alegre Portuguese. A real-time analysis is carried out (LABOV, 1994) and stylistic variation is controlled in the sociolinguistic interviews from which data are collected. The theoretical basis is the Theory of Variation and Change (LABOV, 2008), as well as complementary foundations such as Grammaticalization (GONÇALVES *ET AL.*, 2007), Contextual Styles (LABOV, 2008), and the Decision Tree (LABOV, 2001). Previous studies on this variable (LOPES, 1998; VIANNA AND LOPES, 2012; BORGES, 2004; ZILLES, 2007; SOUZA AND BOTASSINI, 2009; VITÓRIO, 2017) suggest that the variation between the pronouns *nós* and *a gente* is in progress in Brazilian Portuguese. In this investigation, two methodological procedures linked to variationist sociolinguistic are applied: variable rule analysis, which involves collecting speech data from sociolinguistic interviews, coding and submitting them to statistical analysis of logistic regression; and the application of Decision Tree criteria to identify monitored vs. spontaneous speech in interviews and code the data for the Style variable. Such procedures draw upon VARSUL (1990) and LínguaPOA (2015-2019) databases. The statistical analysis took place in the software R, in the RStudio interface, considering samples of 12 interviews for each database (VARSUL and LínguaPOA). The results for the three linguistic variables – *Morphemic Mark*, *Formal Parallelism*, *Style* – and the two social variables – *Gender*, *Age* – point to a significant increase in the realization of *a gente* in the most recent sample (65% in VARSUL database vs. 87% in LínguaPOA database). The variation between the pronouns *nós* and *a gente* is favored by the *Formal Parallelism*, in both samples, and by *Gender*, only in VARSUL database. Regarding stylistic variation, the results of the crossing between *Gender*, *Style* and the dependent variable *nós* showed, in the VARSUL database, a preference among men for using the pronoun *nós*, as well as an increase in the realization by both genders of this variant in the [+monitored] contexts. In the most recent database (LínguaPOA), the hypothesis is that the high percentage of the pronoun *a gente* (87%) and the lack of elementary school informants reduce the influence of stylistic variation even in monitored speech contexts.

Keywords: Nós and a gente. Sociolinguistics. Linguistic Variation. Grammaticalization. Style.

LISTA DE ABREVIACOES

CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
FGV	Fundação Getúlio Vargas
IBGE	Instituto Brasileiro de Geografia e Estatística
IDHM	Índice de Desenvolvimento Humano
LÍNGUAPOA	Acervo de entrevistas sociolinguísticas de Porto Alegre
NURC	Norma Urbana Culta
OBSERVAPOA	Observatório da cidade de Porto Alegre
PB	Português Brasileiro
PE	Português Europeu
PUCRS	Pontifícia Universidade Católica do Rio Grande do Sul
SN	Sintagma Nominal
UFPR	Universidade Federal do Paraná
UFRGS	Universidade Federal do Rio Grande do Sul
UFSC	Universidade Federal de Santa Catarina
VARISUL	Varição Linguística na Região Sul do Brasil

LISTA DE FIGURAS, GRÁFICOS, QUADROS E TABELAS

- Figura 1 Árvore de Decisão
- Figura 2 Índices de IDHM de Porto Alegre
- Figura 3 Planilha de dados *nós* e *a gente* (LínguaPOA)
- Figura 4 Planilha de dados *nós* e *a gente* (VARSQL)
- Figura 5 Janela de transcrição em ELAN de uma entrevista do LínguaPOA
-
- Gráfico 1 Percentuais de uso de *a gente* conforme o ano de nascimento dos falantes, estudo de tendência (dados do NURC e do VARSQL de Porto Alegre)
- Gráfico 2 Percentuais de uso de *nós* e *a gente* em quatro comunidades do Rio Grande do Sul: Flores da Cunha, Panambi, São Borja e Porto Alegre (dados da década de 1990)
- Gráfico 3 Percentuais de uso de *a gente* em amostras socialmente estratificadas (dados coletados entre 1980 e 2000)
- Gráfico 4 Amostra VARSQL – Realização de *nós* por *Gênero* e *Estilo*
- Gráfico 5 Amostra VARSQL – Realização de *nós* e *a gente* por *Faixa Etária* e *Gênero*
- Gráfico 6 Amostra LínguaPOA – Realização de *nós* e *a gente* por *Faixa Etária* e *Gênero*
-
- Quadro 1 Os 12 informantes do LínguaPOA
- Quadro 2 Os 12 informantes do VARSQL
-
- Tabela 1 Fatores sociais significativos para o uso de *a gente* em Porto Alegre, anos 1990.
- Tabela 2 Fatores sociais significativos no uso de *a gente*, estudo de tendência, 1970 e 1990 (dados do NURC e do VARSQL).
-
- Tabela 3 Estimativas dos parâmetros do modelo (de regressão logística, modelo linear geral com efeitos mistos) da realização de *nós* na função de sujeito – VARSQL
- Tabela 4 Estimativas dos parâmetros do modelo (de regressão logística, modelo linear geral com efeitos mistos) da realização de *nós* na função de sujeito – LínguaPOA

SUMÁRIO

1 INTRODUÇÃO	10
2 FUNDAMENTAÇÃO TEÓRICA	13
2.1 Teoria de base	13
2.1.1 A sociolinguística variacionista laboviana	13
2.1.2 A alternância estilística	16
2.2 A variação pronominal no português brasileiro.....	24
2.2.1 O que é gramaticalização?	25
2.2.2 A gramaticalização de <i>a gente</i>	28
3 METODOLOGIA	40
3.1 A comunidade de fala de Porto Alegre	40
3.2 As entrevistas sociolinguísticas do LínguaPOA e VARSUL	42
3.3 Análise de regra variável	43
3.3.1 As amostras LínguaPOA e VARSUL.....	43
3.3.2 Os procedimentos metodológicos	45
3.3.3 As variáveis.....	46
3.3.3.1 Variável-Resposta	46
3.3.3.2 Variáveis previsoras	46
3.3.3.2.1 Variáveis linguísticas	46
3.3.3.2.2 Variáveis extralinguísticas (ou sociais).....	47
3.3.4 A codificação dos dados	48
4 APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS	51
4.1 Análise de regra variável (Resultados)	51
4.1.1 Resultados Estatísticos – VARSUL.....	51
4.1.2 Resultados Estatísticos – LínguaPOA	54
4.2 A variação <i>nós~a gente</i> em tempo real.....	55
5 CONCLUSÃO	58
REFERÊNCIAS	60

1 INTRODUÇÃO

Apesar de a linguística ter alcançado o estatuto de ciência apenas no início do século XX, com o *Curso de Linguística Geral* de Ferdinand de Saussure, e de a sociolinguística ter se instituído como área de estudo autônoma a partir dos anos 1960, muito antes do século XIX já se havia notado que as línguas mudam, levando em consideração uma abordagem que dá conta de que “[...] a coerência do comportamento linguístico, e em particular a regularidade das mudanças fonéticas, podia ser derivada de princípios mais gerais, de preferência psicológicos” (WEINREICH; LABOV; HERZOG, 2006). No entanto, até a primeira metade do século XX, a língua era, predominantemente, tratada sob um viés homogêneo e individual, como sugere Hermann Paul, ao desenvolver a ideia de que “a língua do falante-ouvinte individual encerra a natureza estruturada da língua, a coerência do desempenho falado e a regularidade da mudança” (WEINREICH; LABOV; HERZOG, 2006).

Este trabalho segue a linha de estudos inaugurada na segunda metade do século XX com a Teoria da Variação e Mudança (LABOV, 2008 [1972]), fundamento da sociolinguística variacionista. Nesse sentido, este estudo entende que a língua é constituída de *heterogeneidade ordenada* (WEINREICH; LABOV; HERZOG, 2006): a variação, antes vista como assistemática na língua, passa a ser considerada ordenada, sistêmica e influenciada tanto por fatores linguísticos quanto por fatores extralinguísticos. Seguindo essa proposta, cujo objetivo é analisar a língua no contexto da fala cotidiana e encontrar os registros mais próximos possíveis da *fala vernacular*, busca-se neste trabalho investigar a variação linguística nas formas pronominais de 1ª pessoa do plural, na função de sujeito, e seu condicionamento por fatores linguísticos e fatores extralinguísticos (sociais e estilísticos).

Neste trabalho, se explora o caráter social da língua em uma comunidade de fala específica, na análise de uma variável gramatical e sua alternância estilística dentro das entrevistas sociolinguísticas (i.e. dos estilos contextuais). O objeto de estudo é a variação das formas pronominais de primeira pessoa do plural (*a gente* ou *nós*) no português da comunidade de fala de Porto Alegre (RS). Para tanto, a investigação da variável-resposta se dá com a união de dois procedimentos atrelados à sociolinguística laboviana: a análise de regra variável e a aplicação dos princípios da Árvore de Decisão (LABOV, 2001).

Além da análise de regra variável, em que se examinará a correlação dos fatores linguísticos e dos fatores extralinguísticos com as realizações de *nós* e *a gente*, para que se obtenha o padrão de funcionamento da língua na comunidade de fala em questão, este trabalho

busca investigar a influência dos estilos de fala em entrevistas sociolinguísticas no uso das variantes. A ideia adotada, de Labov (2001), é a de que há alternância de estilos contextuais nas entrevistas e que este aspecto pode influenciar nas escolhas linguísticas. Assim, a variação está inserida em um *continuum* que vai da fala [-monitorada] à fala [+monitorada].

No que diz respeito à questão estilística, serão consideradas as contribuições de Labov (1966), em seu estudo *The Social Stratification of English in New York City*, no qual se apresentam técnicas para isolar os estilos contextuais, até o dispositivo metodológico denominado Árvore de Decisão (LABOV, 2001). Além disso, serão consideradas as ideias de alguns estudiosos(as), como Baugh (2001), Schilling-Estes (2007), Eckert (2001), Valle e Görski (2014), entre outros, que já se debruçaram sobre a temática estilística.

A inclusão da variável estilística na análise se dá, principalmente, pelo entendimento que, assim como nas entrevistas, a alternância estilística é inerente ao uso da língua nas diferentes situações sociais e pode ocorrer por diversos fatores. A construção de *personae* (ECKERT, 2001) é um deles. Outro é a elevação do grau de monitoramento da fala, em virtude do *Paradoxo do Observador*¹ (LABOV, 2008 [1972]), no contexto de entrevistas. Sendo assim, a natureza social e heterogênea da língua, em conjunto com seu viés estilístico, são aspectos pertinentes ao processo investigativo acerca de como se diz e por que se utiliza determinada variante.

Quanto à variação *nós ~ a gente*, serão revisados estudos sobre a gramaticalização de *a gente* e sobre sua alternância com *nós* em diferentes variedades de português brasileiro – Borges (2004), Zilles (2005, 2007), Vitória (2017). Desses estudos, merecerá especial atenção o de Zilles (2007), por conta de também se valer de dados de Porto Alegre, da amostra do VARSUL dos anos 1990, que apontava o progresso no uso de *a gente* na fala da comunidade.

No exame da variação linguística, é importante que se diferencie dois tipos de análise de regra variável: tempo aparente e tempo real. Em relação ao primeiro modelo (tempo aparente), a análise se dá de maneira sincrônica, isto é, em apenas um ponto no tempo, comparando-se as proporções de realização da variável investigada nos diferentes grupos etários controlados. Já no que diz respeito ao segundo modelo (tempo real), temos uma análise com viés diacrônico: compara-se o padrão de variação em diferentes pontos no tempo.

Neste estudo, a análise realizada será em tempo real, do tipo estudo de tendência. Conforme Labov (LABOV, 1994), os dados analisados nesse tipo de estudo não são dos

¹ Labov (2008) assim enuncia o Paradoxo do Observador: “o objetivo da pesquisa linguística na comunidade deve ser descobrir como as pessoas falam quando não estão sendo sistematicamente observadas – no entanto, só podemos obter tais dados por meio da observação sistemática” (LABOV, 2008, p. 244).

mesmos informantes, mas de informantes de mesmo perfil social, considerando-se a estratificação das amostras em questão. Para tal, usamos neste estudo dados de duas amostras: o acervo de entrevistas sociolinguísticas do VARSUL, composto na década de 1990, e o acervo do projeto LínguaPOA, mais recente, que apresenta entrevistas sociolinguísticas coletadas no período de 2015 a 2019.

Com base nisso, tem-se como objetivos deste trabalho (1) verificar se há alteração na proporção total de ocorrência de *nós* e *a gente* no português de Porto Alegre em 25 anos, (2) investigar quais fatores linguísticos e extralinguísticos condicionam o processo de variação das formas pronominais *nós* e *a gente*, na função de sujeito, e (3) com o auxílio da Árvore de Decisão (LABOV, 2001), analisar a influência dos estilos contextuais na variação do objeto investigado. As hipóteses testadas serão as de que: houve um aumento na proporção do uso de *a gente* no intervalo de 25 anos; que fatores como o *Paralelismo Formal* influenciam na escolha por uma outra das variantes concorrentes; e que há um emprego mais frequente da variante *nós* em estilos de fala [+monitorados] nas entrevistas.

O trabalho estrutura-se em quatro capítulos, além desta introdução. O segundo apresenta a fundamentação teórica do trabalho, revisa outros estudos variacionistas acerca da variação *nós~a gente* e também estudos que discutem a questão estilística, em si mesma, mas que direcionam seu olhar às questões desafiadoras da Árvore de Decisão. O terceiro capítulo é o da metodologia: detalha os objetivos da pesquisa, os procedimentos metodológicos da análise estatística de variação linguística incluindo a variável *Estilo*, apresenta os dois bancos de dados – LínguaPOA (2015-2019) e VARSUL (anos 1990) – que forneceram os dados linguísticos utilizados no estudo. No quarto capítulo, é realizada a análise e discussão dos resultados obtidos. Por fim, no quinto capítulo, são feitas considerações finais a respeito dos resultados das análises realizadas e uma reflexão acerca da contribuição deste trabalho para o tema em questão, além de possíveis desenvolvimentos em pesquisas futuras.

2 FUNDAMENTAÇÃO TEÓRICA

Este trabalho toma como suporte a Teoria da Variação e Mudança (LABOV, 2008 [1972], 1994, 2001) em todos os níveis de análise deste estudo, desde os seus estágios iniciais até a interpretação dos resultados encontrados, sejam eles quantitativos ou qualitativos. O presente capítulo divide-se em subseções, as quais abordam, respectivamente, a teoria variacionista laboviana, os aspectos relacionados à variação estilística em estilos contextuais, a revisão de literatura sobre o uso variável nas formas pronominais de primeira pessoa do plural e a respeito do processo de gramaticalização.

2.1 Teoria de base

Esta seção discorre acerca das teorias linguísticas que norteiam a investigação deste estudo: a sociolinguística variacionista laboviana (2.1.1), fundamental para a análise quantitativa da variação *nós~a gente*, que empreenderemos neste trabalho, e a alternância estilística em entrevistas sociolinguísticas (2.1.2), apoio para a anotação das entrevistas de um dos bancos de dados aqui utilizados, o do LínguaPOA (2015-2019), com base em que controlaremos estatisticamente a influência do estilo de fala sobre o uso das variantes *nós* e *a gente*. Com isso, busca-se apresentar a relação entre a variação linguística e a variação estilística, processos inerentes à língua em sua circulação no espaço social.

2.1.1 A sociolinguística variacionista laboviana

Como afirmamos no início deste trabalho, a metodologia inicialmente utilizada nos estudos linguísticos realizados no século XX não levava em consideração a variação linguística, tampouco contemplava a influência dos fatores sociais (extralinguísticos) sobre a língua. A mudança nesse panorama se deu na década de 1960, com o desenvolvimento da Teoria da Variação e Mudança, de Labov (2008), que tem como propósito apresentar, conforme o autor aponta, não uma nova teoria da linguagem, mas um “novo modo de fazer linguística” (LABOV, 2008, p. 298), com base em princípios e fundamentos que dão lugar à variação linguística e admitem os efeitos de fatores sociais como fatos inerentes à língua. Estudos célebres de Labov em que este novo modelo teórico-metodológico é aplicado são o da centralização variável dos ditongos /aw/ e /ay/ em Martha’s Vineyard (LABOV, 1963) e a estratificação social de /r/ em lojas de departamento nova-iorquinas (LABOV, 1966).

Com base em Antoine Meillet (1912), que acreditava que o século XX assistiria ao desenvolvimento da explicação da mudança linguística encaixada na mudança social, Labov

(2008) molda sua abordagem acerca da língua e questiona os pressupostos de Saussure (2000) e Chomsky (1965), grandes nomes nos estudos linguísticos predominantes até os anos 1960, de natureza estruturalista/formalista. Aqui é importante sublinhar que, em sua abordagem, Saussure evoca o conceito de *langue*, no qual aponta que “é a parte social da linguagem... ela não existe fora de um tipo de contrato estabelecido entre os membros da comunidade” (LABOV, 2008, p. 217). A contradição se dá na medida em que os linguistas que trabalham dentro da tradição saussuriana e chomskiana não levam em conta a vida social da linguagem. Nesse sentido, Labov (2008) formula o *paradoxo saussuriano*: o aspecto social da língua, a *langue*, pode ser estudado pela observação de qualquer indivíduo porque a *langue* é compartilhada por todos os membros da comunidade linguística; já a fala, a *parole*, caracterizada por Saussure como individual, só pode ser observada no contexto social.

Em linguística, Chomsky (1957) surge como expoente ao propor a hipótese inatista da linguagem, os conceitos de competência (conhecimento abstrato que se tem da língua) e desempenho linguístico (realização deste conhecimento abstrato). Labov (2008) reconhece todas as contribuições do linguista gerativista no que tange ao refinamento da análise estrutural da língua, mas questiona sua abordagem pautada na ideia de que “o real objeto do estudo linguístico é uma comunidade de fala abstrata, homogênea, em que todo mundo fala igual e aprende a língua instantaneamente” (LABOV, 2008, p. 218). Questiona também a insistência de Chomsky (1965) em dizer que “o dado da linguística não é o enunciado do indivíduo a ser estudado, mas suas intuições acerca da língua” (LABOV, 2008, p. 218).

A partir dessas questões, as quais mostram um modo de “fazer linguística” que mantinha os fatores extralinguísticos apenas no campo das ideias, dando espaço, na prática, apenas para o estudo abstrato da língua, surge a abordagem laboviana. Embora, para Labov, o termo “sociolinguística” se mostre estranhamente enganoso e redundante, na medida em que “a língua é uma forma de comportamento social” (LABOV, 2008, p. 215), o autor é considerado o grande nome da sociolinguística quantitativa ou variacionista. Teoria da Variação é o nome do corpo de ideias defendidas por Labov, fundamentadoras do estudo das correlações entre variáveis linguísticas e variáveis não linguísticas, essas referentes ao contexto social: “o falante, o interlocutor, o público, o ambiente etc.”. (LABOV, 2008 [1972]: 275). Noções-chave dessa Teoria são apresentadas a seguir, iniciando-se por *comunidade de fala*, aspecto fundamental para o entendimento da abordagem laboviana. Para Labov:

A comunidade de fala não é definida por nenhuma concordância marcada pelo uso de elementos linguísticos, mas sim pela participação num conjunto de normas compartilhadas; estas normas podem ser observadas em tipos de comportamento

avaliativo explícito e pela uniformidade de padrões abstratos de variação que são invariantes no tocante a níveis particulares de uso. (LABOV, 2008, p. 150)

Apesar de reconhecer as dificuldades em lidar com dados extraídos da fala, em função de sua aparente assistemática, a variação na fala é, segundo Labov (2008), ordenada na comunidade de fala, relacionada a certos ambientes linguísticos e grupos de usuários. Essas relações dão corpo a ‘normas de uso’ partilhadas pelos falantes da língua. Em outras palavras, toda a língua é heterogênea ou variável em sua realização nas comunidades de fala, mas a variação é sistemática. Orienta-se por normas de uso, tratadas na pesquisa sociolinguística variacionista como ‘regras variáveis’.

Para chegar a essas regras variáveis, o pesquisador precisa enfrentar o desafio de ouvir e gravar grandes quantidades de dados de fala o mais espontânea possível. Mesmo diante da raridade de determinadas variáveis, como as sintáticas, na linguagem vernacular, Labov (2008) sustenta sua posição neste modelo de análise. Além de mostrar que também há problemas no estudo baseado apenas nas intuições do falante, o autor aponta que há “qualquer coisa de trivial” (LABOV, 2008, p. 220) no procedimento de análise linguística, se o olhar do analista se mantém distante da comunidade de fala.

Na comunidade de fala, então, os indivíduos compartilham normas de uso da língua que sustentam a realização sistemática de variáveis linguísticas. Ao sociolinguista variacionista, interessa esclarecer essas normas ou regras variáveis. Para lidar com isso, recorre à análise estatística de regra variável, pois, de acordo com Battisti (2014, p. 14):

A análise de regra variável, de natureza quantitativa, aponta não só o peso relativo dos condicionamentos estruturais e sociais da variação, mas também a tendência de o processo avançar na comunidade, regredir, ou manter-se estável. Os programas computacionais empregados realizam análise logística de regressão, para verificar a interação de diferentes fatores e seus efeitos sobre as realizações variáveis estudadas. [...] A análise de regra variável só pode ser empreendida com uma grande quantidade de dados de fala. Para obtê-los, o pesquisador realiza entrevistas sociolinguísticas: com perguntas sobre temas do cotidiano, procura estabelecer uma conversa com integrantes da comunidade de fala em estudo. Esses são seus informantes, selecionados conforme os estratos sociais controlados na análise. A conversa com cada um dos informantes é gravada para que, posteriormente, o pesquisador possa retomá-la e analisar a fala em suas formas variáveis.

É com esse tipo de procedimento que o pesquisador pode verificar os efeitos internos e externos no funcionamento da língua e atestar, com os resultados de uma análise empírica, o conceito de *heterogeneidade ordenada* proposto por Weinreich, Labov & Herzog (2006 [1968]). São estes efeitos que condicionam a variação e a mudança linguística. Eles podem ser sociais, como *Faixa Etária* e *Gênero*, ou linguísticos, como *Paralelismo Formal* e *Marca*

Morfêmica, abordados neste estudo. Nesse sentido, Labov (2008) advoga pela necessidade de lançar o olhar às mudanças linguísticas em andamento e às relações que fazem com que elas se manifestem. Sendo assim, o objetivo de estudo do linguista é verificar se há relação da variável investigada com os fatores linguísticos e os fatores extralinguísticos controlados, visto que esses atuam em conjunto, provocando a variação sistemática na comunidade de fala, ou seja, a variação e a mudança não ocorrem num vácuo social. Sofrem influência decisiva dos indivíduos e da comunidade. Havendo variação entre escolhas linguísticas, o componente estilístico entra em cena. A este respeito, Labov (2001) também contribui de maneira significativa com o debate.

2.1.2 A alternância estilística

No contexto da variação linguística inerente à língua, insere-se a questão estilística. Essa envolve a escolha, por parte do falante, de uma ou outra variante concorrente conforme um certo registro – formal, informal, etc. – ou atividade de fala – narrar, dar uma opinião, explicar, contar piadas, fofocar, etc. –, assim dando corpo a diferentes *personae* e estilos.

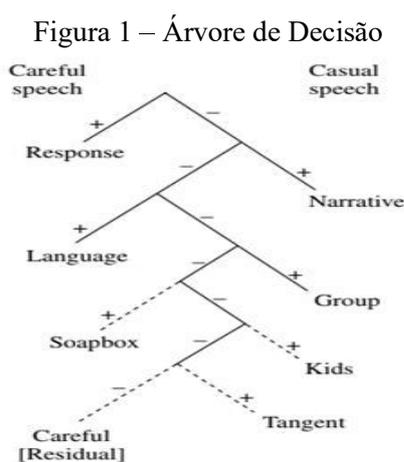
No capítulo *O Isolamento de Estilos Contextuais* (LABOV, 2008), Labov reconhece que a fala do informante na entrevista sociolinguística não é casual como na fala diária, entre pessoas próximas. No entanto, segundo ele, é possível, mesmo na entrevista, obter fala em diferentes estilos, alguns mais espontâneos que outros. Para identificar esses trechos, Labov (2008) apresenta uma maneira de tratar a alternância estilística em entrevistas sociolinguísticas. Para tanto, os trechos das entrevistas sociolinguísticas são segmentados com base em dois eixos: o dos contextos monitorados (de B à D') e o dos contextos espontâneos (de A₁ à A₅), com o pressuposto de que as realizações linguísticas na situação de entrevista estão inseridas em um *continuum unidimensional*, o “da atenção prestada à fala, com a fala casual numa extremidade do *continuum* e os pares mínimos² na outra” (LABOV, 2008, p. 126). Assim, Labov isola estilos contextuais (LABOV, 2008, p. 121-126), procedimento que tem por objetivo segmentar trechos da entrevista sociolinguística e quantificar a atenção prestada à fala.

² Com o objetivo de observar a “variação regular em estilos e contextos diferentes” (LABOV, 2008, p. 101) para as cinco variáveis linguísticas contempladas em seu estudo, o autor controla diferentes estilos de fala do informante na situação de entrevista, os quais chama de ‘estilos contextuais’: A₁ (fala com terceiros ou com o entrevistador antes ou depois da entrevista em si), A₂ (fala com terceiros durante a entrevista), A₃ (fala que não responde diretamente a perguntas do entrevistador), A₄ (parlendas e rimas infantis), A₅ (risco de vida), B (fala em que responde a perguntas formalmente reconhecidas como ‘parte da entrevista’), C (fala na leitura em voz alta de trechos de texto), D (fala na leitura em voz alta de listas de palavras), D' (fala na leitura em voz alta de pares de palavras que se distinguem por um único fonema – os ‘pares mínimos’).

Segundo Bell (2001), a alternância de estilo por parte do falante também sofre a influência de seu interlocutor, isto é, da audiência: busca-se adequar a fala para suprir as expectativas do interlocutor. Outro ponto a ser considerado é o da criação de *personae* e a projeção de posicionamentos via indexação de significados sociais, como visto em Eckert (2000) e Coupland (2007), a respeito da utilização de certos recursos linguísticos com o intuito de desenvolver uma identidade.

A ideia inicial de Labov (1966), embora frutífera como ponto de partida para a análise da alternância estilística, é refinada pelo linguista, juntamente com seus alunos da Universidade da Pensilvânia, com o desenvolvimento do dispositivo que denomina *Árvore de Decisão* (LABOV, 2001). Tal dispositivo é estruturado em dois eixos (fala casual e fala cuidada) e os critérios são dispostos numa escala decrescente de objetividade³. Um esquema representando a *Árvore* está na Figura 1.

Em função do Paradoxo do Observador, isto é, de se buscar nas entrevistas dados de fala casual quando, por sua natureza, tais entrevistas favorecem a produção de fala monitorada, a *Árvore de Decisão* tem como objetivo principal identificar e segmentar trechos espontâneos ou não monitorados de fala do informante, que ocorrem com menor frequência na situação de entrevista. Para tal, a *Árvore* conta com quatro nós contextuais dispostos no eixo da fala cuidada ou monitorada (*careful speech*) — Resposta, Língua, *Soapbox* e Residual — e quatro nós contextuais dispostos no eixo da fala casual ou espontânea (*casual speech*) — Narrativa, Grupo, Infância e Tangente.



Fonte (Labov, 2001, p. 94)

³ “de modo que as primeiras quatro decisões possam ser feitas com o mais alto grau de confiabilidade” (LABOV, 2001, p. 89, tradução nossa).

Começando pelos nós do eixo de fala cuidada, é possível reconhecê-los da seguinte maneira:

Resposta: primeiro trecho de fala que segue a pergunta do entrevistador, seja qual for a produção (por exemplo, uma narrativa, uma opinião, etc.). Corresponde a uma sentença que veicula mais do que um *feedback* ou um eco⁴ da fala do entrevistador.

Língua: trechos que tratam de aspectos linguísticos, incluindo atitudes a respeito da língua.

Soapbox (opinião genérica): opinião de caráter geral, não restrita ao interlocutor-entrevistador, mas expressa como se fosse para uma audiência mais ampla. Fala marcada pela repetição e pela elevação do seu nível de intensidade (volume).

Residual: último contexto estilístico no eixo de fala monitorada da Árvore de Decisão, compreende todos os trechos de fala não classificados nos outros contextos, sejam eles de fala monitorada ou casual.

Já no caso dos nós do eixo de fala casual, a identificação se dá desta forma:

Narrativa: compreende os trechos de fala veiculados através das narrativas orais de experiência pessoal (LABOV; WALETZKY, 2003; LABOV, 1997).

Grupo: fala direcionada a outros interlocutores que, além do entrevistador, porventura estejam presentes na situação de entrevista.

Infância: relatos sobre jogos ou experiências de infância, enunciadas explicitamente em primeira pessoa.

Tangente: fala evasiva, que desvia do último tópico estabelecido pelo entrevistador e é mantida em seu curso somente por interesse do entrevistado.

Percebe-se, com base na literatura (Valle e Görski, 2014; Schilling-Estes, 2007) e por toda a exploração realizada por mim e colegas de LínguaPOA, que, desde o início de uma análise que envolva a Árvore de Decisão, a tarefa de seguir os critérios estabelecidos no dispositivo, até mesmo nos nós mais objetivos, como no nó *Resposta*, é desafiadora. Segundo Valle e Görski (2014), o primeiro percalço se dá na medida em que não é tarefa fácil distinguir o primeiro trecho que segue a fala do entrevistador, de um lado, dos *feedbacks* e ecos, do outro, que não compõem o nó *Resposta* e, portanto, devem ser alocados em outro nó do dispositivo. Nesse sentido, as autoras questionam, especialmente no caso de repetições por parte do

⁴ De acordo com Marcuschi (2003), durante a entrevista, o pesquisador produz sequências voltadas a orientar o falante e monitorá-lo quanto à recepção, isto é, *feedbacks* (Foi tão forte assim?) ou ecos (Como é a tua família? – Como é a minha família?). O mesmo pode ocorrer por parte do informante nas entrevistas em que este está no controle da situação conversacional.

informante da última fala do entrevistador, se “a parte repetida deve ser considerada resposta ou deve ser entendida como *feedback* ou eco? A entonação interfere nessa decisão – se for semelhante, é eco; se for diferente, é resposta?” (VALLE; GÖRSKI, 2014, p. 108).

Outro desafio metodológico desse nó diz respeito aos momentos em que o entrevistador interrompe o curso de fala do informante. No modelo proposto por Labov (2001), não há nenhuma indicação sobre a necessidade de segmentar os trechos de fala do informante que seguem essas interrupções como resposta e, depois, retornar à segmentação referente à atividade de fala anterior, conforme ocorre no exemplo abaixo, de um trecho da entrevista do informante 117 da amostra do LínguaPOA:

D1: *E por que que tu escolheu essa profissão? Foi algo que aconteceu ou tu já pensava isso fazia tempo?*

S1: [na verdade eu posso dizer que foi algo que aconte que eu me descobri sem querer] **então a minha formação ã eu fiz eh como eu sou um cara mais antigo né eu fiz o segundo grau, eu fiz o segundo grau técnico, ã então eu fi eu fiz técnico em eletrônica no colégio Santo Inácio lá onde é a Arena do Grêmio hoje, né, era o colégio, e e aí em seguida fui trabalhar na Varig. eu fazia parte técnica da manutenção dos componentes internos do do dos aviões e odiava só que era coisa de mãe entendeu ah tu tem que ter uma profissão, então e como eu sempre eu fui um cara eh como eu fui sempre um cara estudioso, então assim eu fui ah tá bom vou lá, então assim acabei me formando, ã fui trabalhar porque a Varig na época só selecionava ã alunos profissionais desse colégio, tá eh e e fui trabalhar fazia algo que eu não gostava, trabalhava com algo que eu não gostava, mas trabalhava, e aí quando a Varig entrou em crise lá nos áureos tempos, ã que era uma estratégia vamos dizer assim organizacional totalmente equivocada na época, ela demitiu todos os funcionários que eram os mais novos porque era mais barato independente de ser ã bom profissional ou não era o que era mais barato e eu fui nessa leva, e e aí eu lembro até hoje que eu cheguei em casa ainda morava com meus pais eu tinha acho que dezoito anos dezoito pra dezenove, e aí a minha mãe bah e agora como é que vai ser aí eu disse ah mãe relaxa ganhei um dinheirinho porque eu eu tava dois anos nunca tinha tirado férias então assim aí comprei meu carro com o dinheiro eh e aí eu ti acho que uns vinte e poucos dias depois eu tive um convite de um amigo que tava começando a trabalhar numa filial duma empresa que era de Caxias lá da tua terra que era uma revenda da Microsoft ã e eles iam abrir um escritório aqui e eles precisavam de alguém pra fazer comercial só que eu nunca tinha trabalhado com isso.**

D1: *publicidade assim?*

S1: [não trabalhar com vendas com vendas não tinha salário fixo não tinha carteira assinada não tinha porra nenhuma], e aí eu disse bah eu quero e aí eu me achei [...].⁵

No caso do nó *Narrativa*, primeiro nó do eixo de fala casual e, portanto, o mais objetivo deles, é onde residem os trechos de fala veiculados através das narrativas orais de experiência pessoal. Na situação de entrevista, é o estilo de fala em que há maior possibilidade de emergir a fala casual. No que diz respeito à segmentação desse nó, um dos desafios é a influência do Paradoxo do Observador nas entrevistas sociolinguísticas, isto é, a relação falante-ouvinte na

⁵ Itálico: pergunta do(a) entrevistador(a). Colchete: Trecho referente ao nó resposta. Negrito: Trechos narrativos.

situação formal não é a mesma que se pode captar na fala casual, “nas ruas de Nova York, em bares, no metrô, na praia, ou sempre que visitamos amigos na cidade” (LABOV, 2008, p. 102).

Na medida em que o(a) entrevistador(a) habilidoso(a) consegue criar estratégias para enfraquecer a formalidade da entrevista e superar o Paradoxo do Observador, um novo desafio reside na natureza das narrativas encontradas nas entrevistas sociolinguísticas. Labov (2001), em seu modelo, deixa como indicação que só devem ser segmentadas como narrativas os trechos que veiculam experiências pessoais. Porém, ao lidar com entrevistas sociolinguísticas, o pesquisador frequentemente se depara com narrativas que não se encaixam no padrão indicado por Labov (2001), como as pseudonarrativas (trechos que se iniciam como uma narrativa, mas em que não há complemento) e narrativas vicárias (relatos de experiências vividas por outras pessoas). Nesse sentido, fica a dúvida sobre em qual nó alocar tais trechos, visto que, assim como no caso dos *feedbacks* e *ecos*, mencionados no nó *Resposta*, não há um local adequado para eles.

Com relação aos nós *Língua* e *Grupo*, aquele pertencente ao eixo monitorado e este ao eixo casual, não se apresentam maiores dificuldades em sua aplicação, mas, nas entrevistas de um dos acervos aqui utilizados, as do LínguaPOA (2015-2019), há pouca presença de trechos que podem ser alocados nestes estilos contextuais. A estrutura do roteiro das entrevistas sociolinguísticas utilizadas nas investigações de Labov (2008) difere de forma significativa do roteiro das entrevistas do VARSUL (anos 1990), das analisadas em Valle e Görski (2014) e também do roteiro das entrevistas do LínguaPOA, o que gera um contraste relevante nos resultados. Nas investigações labovianas, havia uma parte do roteiro das entrevistas dedicado à língua, o que não ocorre, por exemplo, na amostra do LínguaPOA, que apresenta uma porcentagem muito incipiente de segmentação do nó *Língua*.

No que diz respeito ao nó *Grupo*, Labov (2001) afirma que momentos anteriores e posteriores à situação de entrevista são eficazes para captar conversas do informante com terceiros, geralmente de fala casual, por isso interessam ao pesquisador. No entanto, assim como no nó *Língua*, não há uma presença significativa de segmentação de trechos neste estilo contextual nas entrevistas dos bancos de dados aqui utilizados. Sendo assim, principalmente em relação ao banco de dados do LínguaPOA, não se verificam muitos dados neste estilo contextual.

Partindo para os nós mais subjetivos, temos como primeiro deles o nó *Soapbox*, situado no eixo monitorado. Labov (2001) indica que os trechos pertencentes a esse nó são os que veiculam uma opinião de caráter geral, não restrita ao interlocutor-entrevistador, mas expressa

como se fosse para uma audiência mais ampla. No entanto, em diversos bancos de dados, como LínguaPOA e VARSUL, são recorrentes as manifestações de opiniões de cunho pessoal, diferentes das narrativas de experiência pessoal. Sem uma indicação no modelo laboviano sobre o que fazer nesses casos, a solução imediata é a de alocar esses trechos no nó *Residual* (explicado adiante), o que acaba sobrecarregando esse nó.

Voltando ao eixo casual, temos o nó *Infância*, onde residem relatos sobre jogos ou experiências de infância, enunciadas explicitamente em primeira pessoa. Aqui, a dificuldade se encontra na medida em que narrativas e o tópico da infância se interseccionam em muitos trechos analisados, o que faz com que seja necessário um olhar mais acurado ao dado de fala. O mais importante passa a ser a perspectiva do informante⁶ em relação a esse relato. Além disso, há trechos em que o informante faz comentários avaliativos sobre sua infância, de forma distanciada, o que gera outro percalço na segmentação desse estilo contextual, uma vez que Labov não faz nenhuma indicação sobre esses casos.

No eixo de fala casual, também temos o nó *Tangente*, que acolhe fala evasiva, aquela que desvia do último tópico estabelecido pelo entrevistador, atendendo a um interesse do entrevistado⁷. Segundo Labov (2001), identificar tais trechos aponta para a necessidade de diferenciar a fala conduzida ora pelo entrevistado, ora pelo entrevistador, isto é, de distinguir quem está no controle da interação.

Por fim, temos o nó *Residual*, onde ficam todos os trechos de fala não classificados nos outros contextos. Por um lado, já se imagina o quão sobrecarregado esse nó pode ficar, na medida em que todas as questões não resolvidas do dispositivo são encaixadas nesse estilo contextual. Por outro, tem-se o problema do eixo de fala em que esse nó está situado. No modelo laboviano, por se tratar de uma situação formal de entrevista, esse nó se encontra no eixo de fala monitorado. No entanto, Labov (2001, p. 93) afirma que “não necessariamente todo o conteúdo residual terá um caráter estilístico comum”, o que indica que se trata de um nó que abarca conteúdos de diferentes naturezas.

Desde que Labov (2001) propôs a Árvore de Decisão e os estilos contextuais, os estudiosos vêm discutindo as vantagens e limites do modelo. Por exemplo, Baugh (2001) propõe uma nova alternativa à configuração binária da Árvore de Decisão (fala casual vs. fala cuidada ou monitorada). No sentido micro, o autor propõe a mudança dos nomes dos dois eixos

⁶ Neste caso, a perspectiva do informante está relacionada ao relato ser contado em primeira ou terceira pessoa. Labov (2001) só considera o trecho como pertencente ao nó infância se o relato for realizado em primeira pessoa, sem que se esteja distanciando da ação.

⁷ Cabe aqui salientar que, se o tangenciamento suscitar outro nó, como uma narrativa, prevalece a hierarquia da Árvore de Decisão (Labov, 2001), ou seja, o trecho seria marcado como narrativa.

estilísticos para fala formal (em vez de fala cuidada) e fala informal (em vez de fala casual). Já no sentido macro, como mudança mais significativa, Baugh (2001) argumenta que cada nó deste novo modelo de Árvore de Decisão aponta para uma tendência de combinação entre os dois estilos em entrevistas sociolinguísticas.

Ainda, segundo John Baugh (2001, p. 115):

Na melhor das hipóteses, quaisquer designações entre fala cuidada e casual deveriam ser especulativas, e os pesquisadores devem ser livres para introduzir evidências relevantes, incluindo dados etnográficos, que forneceriam um número maior de fatos empíricos às análises subsequentes da proporção entre fala cuidada e fala casual, bem como os fatores que podem promover a realização de uma forma em detrimento da outra durante o curso das entrevistas sociolinguísticas.⁸

Já Eckert (2001), em seu artigo *Style and social meaning*, destaca que Labov (2001) não está diretamente preocupado em definir estilo, mas sim em utilizar o estilo para controlar e identificar variação intrafalante. No que diz respeito à Árvore de Decisão, a autora aponta que o objetivo de Labov fora construir um dispositivo “arranjando subeventos da entrevista sociolinguística de acordo com seu potencial em promover variantes padrão e variantes vernaculares” (ECKERT, 2001, p. 119). Ainda, em consideração aos subeventos do dispositivo metodológico laboviano, Eckert (2001) afirma que eles possuem naturezas distintas. A autora aponta que o nó *Resposta* é definido pelo seu lugar no discurso; já o nó *Língua e Infância* têm sua natureza determinada pelo tópico; os nós *Soapbox* e *Narrativa*, em contrapartida, são definidos pelo gênero discursivo; e o nó *Grupo*, pela audiência.

Com relação à atenção prestada à fala, Eckert (2001, p. 122) afirma que:

É importante ressaltar, porém, que a atenção prestada à fala está relacionada a significados sociais heterogêneos: determinados gêneros podem ocorrer em situações que promovem um grau de atenção maior ou menor, certos tópicos podem fazer com que a pessoa se conscientize de como está falando, diferentes interlocutores podem deixar os informantes nervosos em relação a sua autoimagem, e assim por diante.⁹

Nesse sentido, a autora lança luz para a abordagem de que, além da atenção prestada à fala, há outros fatores que dialogam na análise da alternância estilística, como audiência, tópico,

⁸ “Any a priori designations regarding distinctions between Careful and Casual Speech would be speculative at best, and researchers should be free to introduce the relevant evidence, including ethnographic data, that would lend greater empirical verification to ensuing analyses of the ratio of careful to casual speech, as well as the circumstances that may give rise to one form over the other during the course of sociolinguistic interviews” Tradução nossa.

⁹ “It is important to point out, though, that attention paid to speech is related to socially meaningful motivations in heterogeneous ways: certain genres may tend to occur in situations that afford greater or lesser attention to one’s performance, certain topics may remind people of how they are speaking, different interlocutors may make speakers nervous about their self-presentation, and so forth”. Tradução nossa

gênero, construção de *personae*, entre outros. Levando essas questões em consideração, Eckert (2001, p. 122-123) afirma que:

Não há nada que demonstre que a produção da fala vernacular em uma determinada atividade de fala se deve a um grau menor de atenção, e não a outros aspectos. De fato, o audiomonиторamento tende a diminuir quando estamos em um momento relaxado, nos divertindo, quando assumimos o controle do discurso e falamos sobre coisas com as quais nos importamos, etc. Mas não acho que seja claro que a diminuição do audiomonиторamento ou o ato de estarmos sendo espontâneos seja o mais saliente em nossas realizações de variantes vernáculas. Provavelmente, trata-se dos dois fatores e muitos outros.¹⁰

Por fim, fica clara a intenção da autora em propor uma análise multidisciplinar da alternância estilística, levando em consideração o modelo de atenção prestada à fala, proposto por Labov (2001, 2008), mas também aspectos de outras naturezas, como a construção de *personae* e a consequente utilização de determinadas variantes com o intuito de marcar a participação do falante em um grupo social (ECKERT, 2001 p. 124-125), desenvolvendo novas perspectivas na análise de resultados relacionados à variação estilística como um todo.

Schilling-Estes (2007) também discute a *Árvore de Decisão* enquanto modelo de análise. Para a autora, as entrevistas sociolinguísticas são uma fonte valiosa de estilos de fala, bem como de informações sobre as motivações para sua alternância, mas o tratamento da alternância estilística implica ir além das conceitualizações tradicionais e unidimensionais da variação estilística. É preciso inserir, no trato da alternância estilística, abordagens que levem em consideração aspectos etnográficos e interacionais. Para tal e com a ideia de que o “repertório das pessoas no dia a dia compreende uma variedade de estilos autoconscientes e inconscientes” (SCHILLING-ESTES, 2007, p. 971), a autora discorre acerca das críticas direcionadas à abordagem de atenção prestada à fala de Labov (2008), tais como seu foco apenas nessa dimensão, a dificuldade em quantificá-la, assim como a tarefa de identificar, de maneira confiável, os estilos contextuais na *Árvore de Labov* (2001).

Schilling-Estes (2007) propõe agregar aspectos da Sociolinguística Interacional, como *framing* ‘enquadre’, *footing* ‘alinhamento’ e *participation framework* ‘esquema de participação’, propostos por Goffman (1974), e a ideia de que “falantes sempre moldam sua fala de forma a se adequar à situação ou veicular seus propósitos, mesmo que seja sem estar particularmente autoconscientes” (e.g. ECKERT, 2000; SCHILLING-ESTES 2001, 2004), à

¹⁰ “there is nothing to demonstrate that the emergence of the vernacular in a particular speech activity is due to lack of attention, and not other aspects of the activity. Indeed, audio-monitoring is likely to decrease when we’re having fun being ourselves, when we take control of the discourse and talk about the things we care about, etc. But I don’t think it’s clear that it’s the lack of monitoring or the act of being ourselves that is the most salient in our production of vernacular variants. More likely, it is both and a number of other things as well”. Tradução nossa

análise da alternância estilística em entrevistas sociolinguísticas. Schilling-Estes (2007, p. 977-982) ilustra, por meio de trechos de entrevistas, diferenças significativas entre seus resultados e o que poderia se esperar ao seguir a *Árvore de Decisão* como dispositivo metodológico.

Entre seus achados, a autora constata o aspecto performativo das narrativas na entrevista de uma de suas informantes, que a autora chama de Felícia (SCHILLING-ESTES, 2007), o que pode fazer com que esse estilo contextual, considerado pelos variacionistas como “a atividade discursiva mais buscada no trabalho de campo da sociolinguística, uma vez que alegadamente é onde se realiza a fala mais inconsciente e vernacular” (SCHILLING-ESTES, 2007, p. 976), perca esta “fama”, isto é, apresente resultados que não condizem com a “vernacularidade” previamente esperada.

Outro exemplo significativo trazido por Schilling-Estes (2007) é a discussão acerca do nó *Língua* que, *a priori*, possui um caráter formal, o que o coloca como um estilo contextual mais suscetível à realização de variantes padrão. No entanto, a autora mostra (SCHILLING-ESTES, 2007, p. 978) que nem sempre os falantes assumem um elevado grau de atenção prestada à fala ao discorrer sobre a língua. No trecho utilizado por Schilling-Estes (2007), é possível perceber elementos, tais como as pistas do canal, propostos por Labov (2008), assim como o alongamento em expressões vernaculares como “*Whassup?*” (E aí!), o que não é esperado nessas situações. Além disso, a autora menciona a possibilidade de se classificar o trecho no nó *Grupo*, uma vez que há o ingresso de outra pessoa durante a situação conversacional, o que mostra, mais uma vez, a dificuldade de encaixar os trechos nos contextos estilísticos da *Árvore de Decisão*.

Nesse sentido, embora Baugh (2001), Eckert (2001) e Schilling-Estes (2007) abordem desafios à utilização da *Árvore de Decisão*, valorizam-no à medida que propõem ajustar o modelo de Labov (2001, 2008), não descartá-lo. Sinalizam a necessidade de olhar para essa variável indo além da unidimensionalidade. Com isto feito, é possível refinar a análise da alternância estilística e otimizar dispositivos metodológicos, como a *Árvore de Decisão*, ao expandir os cenários nos quais seus princípios possam ser plenamente aplicados.

2.2 A variação pronominal no português brasileiro

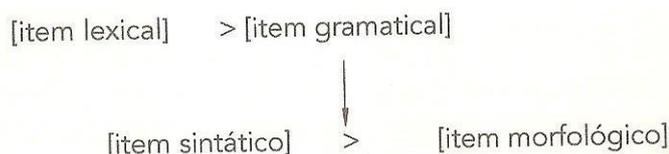
Nesta seção, apresentam-se trabalhos acerca da variação nas formas de expressão de primeira pessoa do plural (*nós* e *a gente*). A primeira subseção (2.2.1) discorre acerca da gramaticalização como um processo. Aborda-se a gramaticalização para esclarecer a mudança na locução *a gente* que passou a funcionar como pronome. A segunda subseção (2.2.2) revisa

os trabalhos de Zilles (2007) e Vitório (2017), que discutem o processo de variação pelo qual passam as variantes pronominais. Busca-se apontar os principais pontos de cada trabalho, desde a metodologia até os resultados, para situar a variável em estudo nesta pesquisa.

2.2.1 O que é gramaticalização?

De acordo com Meillet (1912), configuram-se como gramaticalização¹¹ mudanças linguísticas em que um item lexical¹² autônomo passa a ter uma natureza gramatical, conforme ilustrado no esquema (1) em um sistema de *clines* de mudança, com a passagem do [léxico] > [gramática], sendo este último composto pela sequência [sintaxe] > [morfologia]:

(1)



Fonte: Gonçalves *et al.* (2007, p. 22)

Ainda, segundo Kurilowicz (1975 [1965]), a gramaticalização é um processo de morfologização, que faz com que um item passe de uma natureza lexical para gramatical, mas também opera sobre itens que já possuem esta natureza, aumentando ainda mais o seu nível de gramaticalidade. Tal proposta é observável no esquema (2), fornecido por Hopper e Traugott (1993) que, assim como Meillet (1912), admitem que uma forma em processo de gramaticalização segue um *cline* de mudança:

(2) [item de conteúdo] > [palavra gramatical] > [clítico] > [afixo flexional]

Fonte: Hopper e Traugott (1993, p. 7)

Apesar do primeiro uso do termo ser atribuído a Meillet (1912), é válido salientar que antes disto já era possível identificar estudos de gramaticalização. Tais pesquisas datam dos séculos X, na China, século XVII, na França (Condillac e Rosseau) e Inglaterra (Tooke), e

¹¹ Apesar de haver controvérsias acerca do nome desse processo (GONÇALVES *et al.*, 2007, p. 16-17), a terminologia adotada neste estudo será gramaticalização. Além disso, apesar da possibilidade de se analisar a gramaticalização em uma perspectiva de [qualquer material linguístico] > [+gramatical], este estudo será pautado na combinação entre as concepções de Meillet (1912) [lexical] > [gramatical] e Kurilowicz (1975) [-gramatical] > [+gramatical].

¹² Tem-se como ideia de item lexical categorias prototípicas cujas propriedades fazem referência a dados do universo bio-psíquico-social, designando entidades, ações, processos, estados e qualidades. (GONÇALVES *et al.*, 2007, p.17)

século XVIII, na Alemanha (Bopp, Schlegel, Humboldt e Gabelentz) e Estados Unidos (Whitney). A partir do século XX, é possível destacar as pesquisas de linguistas da Alemanha (Lehmann, Heine, Claudi, Hünemeyer) e da Costa Oeste Americana (Gívon, Hopper, Traugott, Bybee, Pagliuca, entre outros).

Passado este breve percurso histórico e conceitual sobre o que é gramaticalização, busca-se, aqui, apresentá-la como processo que envolve variação linguística. Zilles (2007, p. 28) aponta a gramaticalização como “um tipo particular de mudança linguística em virtude de certas características que lhe são inerentes”.

De acordo com a autora, tem-se na gramaticalização:

um processo **linguisticamente motivado** e altamente encaixado no sistema linguístico. Assim, os fatores decisivos que lhe dão início e continuidade não estão exclusivamente nos itens que se gramaticalizam, mas em mudanças simultâneas, que estão ocorrendo ou já ocorreram em subsistemas linguísticos relacionados. Fala-se, então, em **feixe de mudanças inter-relacionadas** (ZILLES, 2007, p. 28. Grifo da autora).

Para ilustrar o conceito de mudanças inter-relacionadas, Zilles (2007) recorre ao processo de pronominalização, como o de *você* no português, que modifica a configuração do paradigma da concordância verbal da língua portuguesa. Outra característica atrelada ao processo de gramaticalização é o seu curso. No que diz respeito a esse aspecto, Zilles afirma que “como as demais mudanças linguísticas, ela ocorre por transições graduais ao longo de um *continuum* que se projeta no tempo” (ZILLES, 2007, p. 29), isto é, um processo unidirecional.

Com relação ao princípio de unidirecionalidade da gramaticalização, é necessário que se diga que não há um consenso sobre o tema. Alguns autores, como Craig (1991), apontam que ocorre uma “poligramaticalização” na gramaticalização, isto é, a ideia de que uma forma simples passa por diversos caminhos de mudança até alcançar seu estatuto gramatical. Castilho (2002a), por exemplo, sugere que se abandone o princípio da unidirecionalidade em prol da multidirecionalidade, propondo que enxerguemos a língua como um multissistema dinâmico composto por quatro domínios: léxico, semântica, discurso e gramática.

Já Hopper e Traugott (1993, p. 115), em relação ao conceito de poligramaticalização de Craig (1991), apontam que o “desenvolvimento para domínios gramaticais diferentes, via caminhos tão múltiplos, conforma-se à unidirecionalidade, uma vez que formas resultantes são mais gramaticais (abstratas, reduzidas, generalizadas) do que as formas originais”.

Além disso, Gonçalves *et al.* (2007, p. 60) sublinham que, ao se posicionarem como contrários ao princípio da unidirecionalidade, os estudiosos cometem “equivocos como

envolver somente itens lexicais na análise, assim como constituir casos claros de lexicalização”. Ainda, atesta-se a necessidade “de um melhor entendimento das relações entre o léxico e a gramática, mesmo entre os partidários da gramaticalização” (Gonçalves *et al.*, 2007, p. 60), na medida em que, segundo os autores, detratores da unidirecionalidade, como Newmeyer (2001), descartam o nível acima da palavra em suas considerações sobre o tema.

Por fim, é importante salientar que as principais críticas ao princípio da unidirecionalidade residem no entendimento de que há uma irreversibilidade conjugada a ele.

Neste sentido, Caetano (2011, p. 64) alerta:

A unidirecionalidade é o princípio pelo qual certos itens (lexicais ou menos gramaticais) tendem, num *continuum*, a deslizar a itens gramaticais ou mais gramaticais. No entanto, não se deve confundir unidirecionalidade com uma suposta irreversibilidade, pois, muitas vezes, ocorre mesmo sincretismo do item em tela, que opera tanto no nível lexical ou menos gramatical, em dados contextos, quanto num nível gramatical ou mais gramatical, em outros.

Para além do princípio da unidirecionalidade, o qual será adotado neste estudo, há outro aspecto importante a ser salientado, que são os mecanismos da gramaticalização. De acordo com Heine (2003), o processo de gramaticalização consiste em um conjunto de mudanças em que se compreendem quatro mecanismos:

- a) Dessemantização (ou “bleaching”, redução semântica): perda de conteúdo semântico;
- b) Extensão (ou generalização contextual): uso em novos contextos;
- c) Decategorização: perda de propriedades morfossintáticas características das formas-fonte, incluindo a perda do status de palavra independente própria da cliticização e da afixação;
- d) Erosão (ou “redução fonética”): perda de substância fonética.

Com relação a estes quatro mecanismos da gramaticalização, Zilles aponta que:

cada uma dessas mudanças pode ser analisada em termos de seu encaixamento linguístico e social, sendo admissível que se tenha diferentes grupos sociais como líderes, que progrida em diferentes ritmos e que receba diferentes avaliações sociais em diferentes contextos (ZILLES, 2007, p. 30).

Dadas essas considerações acerca da gramaticalização como processo, é chegada a hora de revisar as contribuições de Zilles (2007) e Vitória (2017) sobre o percurso da variante *a gente* no português brasileiro. Com isto, busca-se, aqui, explorar a origem de inserção desta forma gramaticalizada na língua portuguesa, assim como o que isto acarreta na rotina de

componentes linguísticos relacionados e o que motiva a preferência por esta nova variante em detrimento da já consagrada variante *nós*.

2.2.2 A gramaticalização de *a gente*

Com evidências em todo o território nacional, o acréscimo do uso da forma *a gente* na variável pronome pessoal de primeira pessoa do plural tem seu destaque entre as reconfigurações pelas quais passa o português do Brasil (ZILLES, 2005, 2007, VITÓRIO, 2017), sendo analisada em estudos quantitativos de análise de regra variável, com base nos pressupostos teórico-metodológicos da sociolinguística variacionista laboviana (LABOV, 2008 [1972]).

Zilles (2007) aponta que, assim como a introdução de *você/vocês* no PB, a inserção de *a gente* na língua portuguesa tem provocado uma redução no paradigma da concordância verbal, sendo mais recorrente uma combinação entre a forma *a gente* + verbo na 3ª pessoa do singular. No entanto, a autora salienta que há registros de combinação entre *a gente* + verbo na 1ª pessoa do plural, embora de maneira incipiente se comparado às ocorrências com verbos na 3ª pessoa do singular (ZILLES; MAYA; SILVA, 2002).

No que diz respeito a outros aspectos envolvendo o encaixamento linguístico de *a gente*, a autora destaca dois deles: o parâmetro do sujeito nulo e do objeto nulo. A respeito do primeiro, Simões (2006) destaca que, na comparação com o PE, o PB é mais restritivo ao emprego de sujeitos nulos, o que aponta para um progressivo preenchimento do sujeito na variedade de Português Brasileiro. Em relação ao objeto nulo, Zilles (2007, p. 31) argumenta que se trata de “uma inovação sintática do PB em relação ao PE que reforça a necessidade de explicitar o sujeito, se não pela desinência, pelo uso de pronome ou SN”. Tais reconfigurações percebidas no sistema linguístico sob a influência da inserção da variante *a gente* corroboram a ideia do *feixe de mudanças inter-relacionadas*, isto é, que os elementos não se movimentam isoladamente na língua e a chegada de uma nova forma acarreta uma reconfiguração de outras partes do sistema.

Em relação à origem da gramaticalização de *a gente*, não há surpresa a respeito do item-fonte do processo, se verificada a literatura. De acordo com Heine e Kuteva (2002, p. 232-233), é comum que nomes genéricos como *homem* e *povo* sirvam como item-fonte em processos de gramaticalização. Sendo assim, é natural que um item lexical como *gente* faça parte desse processo. Conforme Lopes (2001, p. 140-141, *apud* ZILLES, 2007, p. 31), a gramaticalização de *a gente* começa com seu uso como pronome indefinido, expressão indeterminada com

sentido genérico, no século XVI, em substituição ao uso de *homem~ome*. Zilles (2007) aponta que não há clareza sobre a razão dessa substituição, mas que se trata do ponto final de um processo de gramaticalização e o início de um novo ciclo com a promoção de *a gente*.

Além disso, a autora sinaliza que, para a compreensão do fenômeno de gramaticalização de *a gente*, é necessário que se tome contato com características intrínsecas do item lexical fonte. A respeito disto, Zilles cita o fato de que (1) o sentido original do substantivo latino *gens, gentis*, por ser *povo*, tinha, inerentemente, um traço semântico de pluralidade e um traço de pessoa (ser humano) e (2) o substantivo *gente* ser um nome coletivo o favoreceria, em uma próxima etapa do processo, a assumir função de pronome indefinido, com o valor [+genérico] atrelado ao significado de “toda e qualquer pessoa” (ZILLES, 2007, p. 31-32).

No que diz respeito ao conjunto de mudanças exposto por Heine (2003), Zilles (2007) discute de maneira didática como cada um deles se movimenta no processo de gramaticalização de *a gente*. Em relação à *dessemantização*, a autora afirma que o item-fonte *gente* perde o seu traço de povo, porém mantém o de pessoa. Ainda, Zilles (2007) aponta que este aspecto é fundamental na próxima etapa do processo, “quando ocorre a posterior mudança semântica relacionada com a possibilidade de *a gente* expressar pessoa do discurso: eu e tu, eu e outras pessoas” (BORGES, 2004). No segundo mecanismo do processo, a *extensão*, Zilles (2005) afirma que o uso de *a gente* na função de sujeito aumenta de maneira significativa entre as décadas de 1970 e 1990.

A este respeito, a autora aponta que:

O momento dessa aparente aceleração da mudança coincide com uma transformação social substancial no Brasil afetando a demografia, a geografia e a estrutura socioeconômica do país. Esta transformação envolveu a industrialização, a migração das áreas rurais para áreas urbanas, um enorme desenvolvimento tecnológico na comunicação, assim como de uma classe trabalhadora urbana, o aumento progressivo de matrículas em escolas públicas, entre muitos outros aspectos (ZILLES, 2005, p. 30)

Além disso, a forma *a gente* passa a ser inserida em novos contextos, do uso genérico (3) para funções de referência específica (4) e pronome anafórico (5) (ZILLES, 2007, p. 32), conforme exemplos abaixo:

(3) No excerto de entrevista a seguir, a informante está falando de outros lugares que conhece além de Porto Alegre: Conheço Bento, né? Rio Grande do Sul, né? até que eu não, não conheço, não conheço muito, né? Eu procuro, estou procurando fazer assim mais pro Rio

Grande do Sul também, né? Que eu não, né? o estado a gente tem que conhecer, né? Tem tanta coisa. (RSPOA31, VARSUL, 1482 a 489).

(4) No excerto de entrevista a seguir, a informante está falando sobre as brincadeiras da infância: (hes) nós não tínhamos muito contato com outras crianças, a vó nunca deixou a gente sair assim [do] – do portão pra fora pra brincar com outras crianças, então o nosso mundo éramos nós mesmos. (RSPOA28, VARSUL, 1.33-37).

(5) a gente, olha pra gente, e vê que... (uso anafórico)

Fonte: Zilles (2007, p. 32-33)

Sobre a *deategorização*, terceiro mecanismo do processo de gramaticalização, Zilles (2007) aponta que, ao se tornar um pronome indefinido, ocorre a fixação da sequência *a gente* e restrições combinatórias, isto é, nas palavras da autora, casos como *a boa gente* não se referindo à função de pronome, mas substantivo; e sequências como *A gente, gaúchos, temos muito orgulho* não sendo uma possibilidade nem com sentido genérico, nem com o sentido pessoal. Em segundo lugar, conforme Carvalho (2015), ocorre a perda do plural gramatical e do gênero feminino fixo, havendo uma concordância de acordo com o referente – o emissor da mensagem –. Em relação ao plural gramatical, segundo a autora, não é possível uma construção como “as gentes” para designar primeira pessoa do plural, como se observa no exemplo extraído de Fischer (2004):

Board – Palavra que tem sido muito usada hoje em dia, tempo de globalização dos mercados, o Terceiro Mundo entregando de bandeja o patrimônio nacional para a grana internacional, essas coisas, e usada no sentido inglês mesmo, de conselho, direção, as gentes que mandam em uma empresa, os chefões. (FISCHER, 2004, p. 44).

Já a respeito do gênero feminino fixo, Zilles (2007, p. 33) ilustra a questão com o seguinte trecho:

(6) *a gente* está muito cínico, assim muito na retaguarda (falante do gênero feminino, referindo-se a quem num determinado partido político) (RSPOA31, 1. 900).

Por fim, sobre o último mecanismo da gramaticalização, a erosão, Zilles (2007) afirma que a variante *a gente* também pode ser realizada como *ahente*, *a'ente* e *'nte*, como se observa em (7) e (8):

- (7) amanhã a'ente vai lá
 (8) depois 'ente pega o ônibus e vai

Fonte: Zilles (2007, p. 33)

A redução fonética de *a gente* não será aprofundada, uma vez que não foi observada nesta pesquisa, mas é digno de nota que registros como os encontrados por Zilles (2007) sugerem que a gramaticalização de *a gente* no português brasileiro se encontra em um estágio bem avançado.

No que tange à etapa da análise de regra variável, Zilles (2007) apresenta resultados quantitativos que ilustram o panorama da variante sincronicamente (tempo aparente) e diacronicamente (tempo real).

Em tempo aparente (tabela 1), a autora lida com 39 informantes do banco de dados do VARSUL, todos de Porto Alegre e estratificados em sexo binário (19 do sexo masculino e 20 do sexo feminino), em duas faixas etárias (17 de 25 a 49 anos e 22 acima de 50 anos) e nível de escolaridade (10 com escolaridade elementar – até 5 anos, 8 com escolaridade intermediária – 5 a 8 anos, 9 com escolaridade secundária – 9 a 11 anos, e 12 com escolaridade pós-secundária – mais de 11 anos).¹³

Tabela 1 – Fatores sociais significativos para o uso de *a gente* em Porto Alegre, anos 1990.

	N/Total	%	Peso
Gênero			
Masculino	422/678	62	0,41
Feminino	915/1266	72	<u>0,55</u>
Idade			
25-50	480/618	78	<u>0,66</u>
50-70	857/1326	65	0,42
Totais	1337/1944	69	Input: 0,85

Fonte: Zilles (2007, p. 34)

Os resultados desse estudo sincrônico sugerem, nas palavras de Zilles (2007), que se trata de uma mudança liderada por mulheres (peso 0,55 vs. peso 0,41 masculino), o que converge com o esperado, de acordo com a literatura (LABOV, 2008). Além disso, a autora afirma que os resultados da variável *Faixa Etária* sinalizam uma mudança em curso, uma vez que os jovens apresentam um peso relativo muito maior do que a faixa etária mais velha (0,66 vs. 0,42). Ainda, Zilles (2007) sublinha que o valor do *input* é muito alto (0,85), indicando uma mudança bastante avançada.

¹³ Segundo Zilles (2007), tais critérios são estabelecidos pela estratificação do VARSUL.

Em relação ao seu estudo diacrônico (tabela 2), também em Porto Alegre, Zilles (2007) lida com duas amostras gravadas num intervalo de 20 anos: a do Projeto NURC (1970) e a do projeto VARSUL (1990). Composta por 36 entrevistas, a estratificação da análise se dá com a presença de 18 informantes do gênero masculino e 18 do gênero feminino, divididos em duas faixas etárias: a dos jovens (25-44 anos) e a dos velhos (45-69 anos). Sobre o grau de escolaridade dos informantes, Zilles (2007, p. 35) alerta que grande parte deles (32) tem educação universitária completa, porém, na amostra dos anos 1990, foram incluídos quatro falantes com educação secundária, supondo não haver diferença significativa entre eles, com base nos resultados do estudo em tempo aparente (ZILLES, 2005). A autora afirma que 20 dos informantes entre os 36 foram ouvidos na década de 1970 e os outros 16 na década de 1990.

Tabela 2 – Fatores sociais significativos no uso de *a gente*, estudo de tendência, 1970 e 1990 (dados do NURC e do VARSUL).

	N/Total	%	Peso
Faixa etária			
Jovens	633/865	73	<u>0,64</u>
Velhos	358/668	54	<u>0,32</u>
Década/entrevista			
1970 (NURC)	403/721	56	0,31
1990 (VARSUL)	588/812	72	<u>0,67</u>
Gênero			
Masculino	387/654	59	0,46
Feminino	604/879	69	<u>0,53</u>
Totais	991/1533	65	Input: 0,74

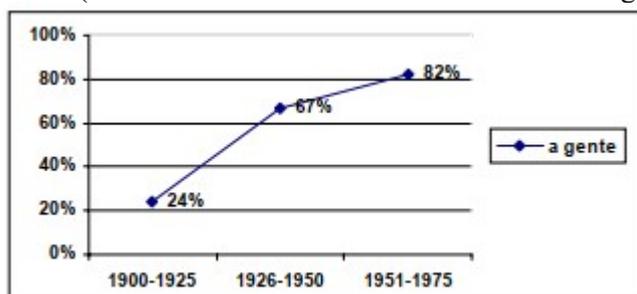
Fonte: Zilles (2007, p. 35)

Os resultados quantitativos encontrados na análise em tempo real ilustram um aumento significativo do uso de *a gente* entre as décadas de 1970 e 1990. Entre os achados, estão uma diferença grande nos percentuais e nos pesos relativos, entre as faixas etárias (mais jovens 0,64 vs. 0,32 mais velhos) e as décadas (0,67 nos anos 1990 vs. 0,31 nos anos 1970). Tal aceleração do processo, nas palavras de Zilles (2007), está atrelada principalmente ao aumento do uso da variante *a gente* pela faixa etária mais jovem. Estes resultados corroboram com o exposto em Zilles (2005) e Borges (2004), este já identificando um aumento na preferência por *a gente* na década de 1960. Ainda, nota-se que a variante *a gente* não sofre estigma¹⁴, visto que a amostra é composta por maioria de informantes com grau de escolaridade bastante elevado.

¹⁴ Na língua falada, um contexto em que o pronome *a gente* costuma sofrer estigma é quando se espera a utilização da primeira pessoa do singular na sentença e o informante opta pela variante gramaticalizada: usa-se *a gente* em detrimento do esperado *eu*.

Zilles (2007) também observa um acelerado avanço no uso de *a gente* em detrimento de *nós*, com relação ao ano de nascimento dos falantes. Tal crescimento pode ser percebido no Gráfico 1 exposto pela autora:

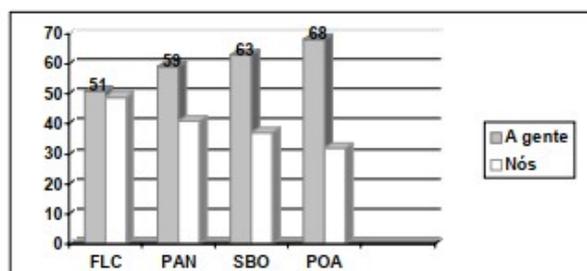
Gráfico 1 – Percentuais de uso de *a gente* conforme o ano de nascimento dos falantes, estudo de tendência (dados do NURC e do VARSUL de Porto Alegre)



Fonte: Zilles (2007, p. 36)

Além disso, Zilles (2007) apresenta resultados interessantes no que diz respeito ao desenvolvimento da realização da variante *a gente* em comunidades bilíngues e em outras partes do território nacional. Em relação às comunidades bilíngues, a autora lança luz sobre Flores da Cunha, em que há contato entre o italiano e o português, e Panambi, onde ocorre o uso do alemão em paralelo ao português. Ainda, a autora cita o caso de São Borja, cidade situada na fronteira com a Argentina, portanto, com uma proximidade muito forte entre espanhol e português. Na comparação com a capital Porto Alegre, o Gráfico 2 mostra uma porcentagem menor do uso de *a gente* nas comunidades bilíngues, mas os quatro locais seguem em direção a um aumento destes valores.

Gráfico 2 – Percentuais de uso de *nós* e *a gente* em quatro comunidades do Rio Grande do Sul: Flores da Cunha, Panambi, São Borja e Porto Alegre (dados da década de 1990)

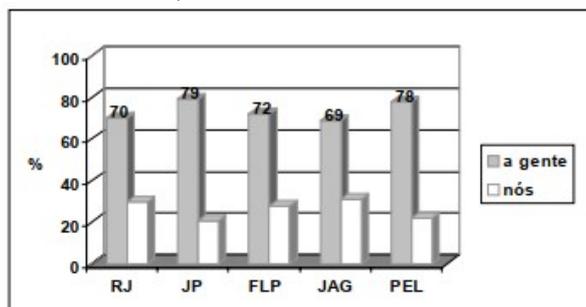


Fonte: Zilles (2007, p. 36)

Em relação ao que se observa nacionalmente, Zilles (2007) compara os resultados de três capitais – Rio de Janeiro, João Pessoa e Florianópolis, com os achados mais recentes (de 2000 em diante) em duas comunidades do Rio Grande do Sul – Pelotas e Jaguarão, que fazem

fronteira com o Uruguai. Assim como no Gráfico 2, se observa um ritmo mais acelerado do uso de *a gente* nas capitais, enquanto comunidades do interior/bilíngues seguem esta mudança num ritmo mais lento.

Gráfico 3 – Percentuais de uso de *a gente* em amostras socialmente estratificadas (dados coletados entre 1980 e 2000)



Legenda: RJ: Rio de Janeiro, Sudeste (OMENA; BRAGA, 1996); JP: João Pessoa, Nordeste; corpus do projeto VALPB (FERNANDES, 1999); FLP: Florianópolis, Sul; corpus do projeto VARSUL (SEARA, 2000); JAG: Jaguarão, Sul; corpus do projeto BDS-Pampa (BORGES, 2004); PEL: Pelotas, Sul; corpus do projeto VarX (BORGES, 2004). Fonte: Zilles (2007, p. 37)

Por fim, Zilles (2007) alerta que, por mais que os resultados apresentados beírem a casa dos 80% para a realização de *a gente* e, portanto, sugiram que não há estigma sob esta nova variante, é preciso ter cautela com a avaliação social desta nova forma pronominal em certos gêneros textuais e, em alguns casos, até mesmo na fala (ZILLES, 2007, p. 38).

Na medida em que se apresenta como uma variante fortemente atrelada à língua falada, à informalidade, e que os resultados quantitativos, como os encontrados em Zilles (2005, 2007), sugiram um maior número de características da variante *a gente* atreladas a uma *mudança vinda de baixo*¹⁵ em vez de uma *mudança vinda de cima*¹⁶ (Labov, 1994), a restrição ao uso desta forma pronominal em textos escritos formais é previsível e muito mais forte na comparação com a sua presença em contextos de fala [-monitorados] e até mesmo em alguns contextos [+monitorados], como em debates, em que já é comum a escolha da forma gramaticalizada por parte dos(a) candidatos(as).

Além das contribuições de Zilles (2007), este trabalho também considera os achados de Vitória (2017) acerca da variação entre *nós* e *a gente*. Neste último, a autora investiga a realização dos pronomes *nós* e *a gente* na função de sujeito e de complemento e adjunto na cidade de Maceió/AL. Para realizar a análise quantitativa (em tempo aparente), a autora utiliza

¹⁵ Para Labov (1994), em linhas gerais, mudanças vindas de baixo (*change from below*) são mudanças sistemáticas que aparecem primeiro no vernáculo e representam a operação de fatores linguísticos internos.

¹⁶ Já no caso das mudanças vindas de cima (*change from above*), Labov (1994) considera aquelas introduzidas pelas classes dominantes, muitas vezes com plena consciência da comunidade.

o programa computacional GoldVarb X, em uma amostra composta por 72 entrevistas de falantes maceioenses, que datam do ano de 2010. Em relação à estratificação, a amostra segue o modelo binário de sexo/gênero – homens e mulheres –, três faixas etárias classificadas como F1 (15-29 anos), F2 (30-44 anos) e F3 (acima de 44 anos), bem como a divisão em três níveis de escolaridade (fundamental, médio e superior).

De acordo com a literatura (OMENA, 2003; LOPES, 1998, 2004; SEARA, 2002; FERNANDES, 2004; ZILLES, 2007; MAIA, 2009; TAMANINE, 2010; FRANCESCHINI, 2011; SANTOS, 2014; RUBIO, 2014; VITÓRIO, 2015, entre outros), Vitória (2017) nos diz que, na função de sujeito, *a gente* é a forma favorita entre as variedades brasileiras e tem como condicionadores linguísticos as variáveis *Paralelismo Formal* e discursivo, a determinação do referente, o tempo verbal, a saliência fônica e o preenchimento do sujeito. Nos condicionadores sociais, se manifestam faixa etária, sexo/gênero e escolaridade. Ainda, a autora aponta que a variante é favorecida pelos fatores (1) menor diferença fônica entre as formas verbais; (2) com traço do referente [+indeterminado]; (3) em formas verbais menos marcadas; (4) quando o verbo que o acompanha está na terceira pessoa do singular; (5) com sujeito pronominal preenchido; (6) pelas mulheres; (7) falantes com nível de escolaridade mais baixo; e (8) entre os falantes mais jovens, o que sinaliza uma mudança em progresso.

Em sua análise, Vitória (2017) estabelece como pressupostos para os resultados da variação entre *nós* e *a gente*, na função de sujeito, o fato da variante *a gente* ser a preferida entre as variedades brasileiras, bem como os condicionadores supracitados. Os valores resultantes da análise e rodada de dados da pesquisa de Vitória (2017) mostram 624 realizações das formas em competição na posição de sujeito na fala maceioense, sendo 100 destas realizações referentes ao pronome *nós* e as demais 524 ao pronome *a gente*. Tais valores representam um percentual elevado de realização do pronome *a gente* (84%) em comparação com seu concorrente *nós* (16%), o que confirma a hipótese de trabalho da autora e corrobora com os resultados de Zilles (2007), a respeito da realização da forma *a gente* nas capitais brasileiras.

Entre os grupos de fatores potencialmente relevantes destacados pelo programa de análise estatística usado pela autora, o GoldVarb X (*Marca Morfêmica, Paralelismo Formal, Preenchimento do Sujeito, Escolaridade, Faixa Etária*), Vitória (2017) inicia sua exposição pela *Marca Morfêmica*. Este primeiro condicionador diz respeito à concordância verbal com os pronomes *nós* e *a gente*, isto é, o verbo que acompanha as variantes. Com base nos dados da análise de Vitória (2017), nota-se a possibilidade de duas combinações para *nós* e *a gente*, no que diz respeito a sua concordância verbal: *nós* e *a gente* + P4 (morfema -mos), nos exemplos

(9) e (10) e *nós* e *a gente* + P3 (morfema zero), nos exemplos (11) e (12). Conforme visto em Zilles (2007), a autora assume como pressuposto o favorecimento de morfema zero pela variante gramaticalizada, uma vez que *a gente* tende a manter seu traço de terceira pessoa.

(9) - é de dois em dois anos *nós fazemos* reciclagem desse curso. (L12L1781)

(10) eu acho que *a gente* hoje *temos* o melhor prefeito que tem na história de Maceió. (L3L363)

(11) - com três tiros na cabeça é o que *nós* até agora *ficou* sabendo. (L9L1284)

(12) *a gente* faz todo o serviço burocrático – atendimento ao público. (L53L4688)

Fonte: Vitório (2017, p. 126)

Os valores encontrados (99% de P3 + *a gente* vs. 24% de P4 + *a gente*) por Vitório (2017) dão conta de que o morfema zero, de fato, favorece a realização de *a gente*, o que corrobora com Zilles (2007). Em relação aos pesos relativos (-mos 0.3 vs. zero .81), a autora aponta que os resultados percentuais confirmam a hipótese de que o morfema *-mos* inibe a realização de *a gente*, enquanto o morfema zero favorece, convergindo com o entendimento de que *a gente* tende a manter o traço formal e original de terceira pessoa concernente ao nome coletivo *gente*.

Nos resultados do segundo fator considerado na análise, o *Paralelismo Formal*, que, segundo Omena (1996, *apud* VITÓRIO, 2017 p. 126) é entendido como a tendência de o falante repetir uma mesma forma em uma dada sequência discursiva, foram observadas quatro possibilidades de combinação para as variantes *nós* e *a gente*, sendo elas *isolada*, conforme exemplo (13), primeira *da série*, no exemplo (14), assim como *antecedido por nós* e *antecedido por a gente*, nos exemplos (15) e (16), respectivamente. A autora assumiu como pressuposto o favorecimento do paralelismo antecedido por *a gente* na realização da variante *a gente*.

(13) porque *a gente* tá vendo que ele tá trabalhando – você anda em qualquer canto de Maceió você vê obras – e Maceió agora tá diferente. (L12L1707)

(14) - tá entendendo – num tem essa segurança – às vezes *a gente* quer trabalhar até mais tarde um pouquinho Ø num consegue – Ø num pode trabalhar – você não pode – eu tô aqui sozinha /mais, mas/ eu tô ligada. (L20L2878)

(15) eu trabalho num hospital que foi agora municipalizado e *nós* não temos um comprimido pra dor – tudo que *nós* temos lá é injetável e isso é um absurdo. (L14L1980)

(16) *a gente* merecia muito mais *a gente* num tem um policiamento aqui de dia e de noite *a gente* tem um certo receio por conta de que *a gente* tem medo o índice de ladrão aqui é muito grande. (L14L2060)

Fonte: Vitório (2017, p. 127)

Assim como os resultados encontrados no fator *Marca Morfêmica*, a hipótese estabelecida pela autora se confirma, uma vez que a forma *antecedida por a gente* aparece com um percentual de 97% e um peso relativo de .82, em contraste com a forma *antecedido por nós*, que apresenta um percentual de 21% e um peso relativo de .02.

Em relação às formas *isolada* e *primeira da série*, os resultados mostram que ambas desfavorecem a realização da variante *a gente*, apresentando um percentual de 84% e peso relativo .26 e um percentual de 87% e um peso relativo de .27, respectivamente. Tais resultados convergem com outros estudos, como o de Vitório (2015), que apontam que a forma *a gente* tende a ser realizada com mais frequência em situações em que é precedida por si mesma, num contexto de *Paralelismo Formal*.

A terceira variável linguística selecionada na variação de *nós* e *a gente*, na função de sujeito na fala maceioense – o Preenchimento do Sujeito –, refere-se à expressão plena ou nula do sujeito pronominal. A este respeito, Vitório (2017) indica que expressão plena são os casos em que as formas pronominais *nós* e *a gente* são expressas foneticamente na indicação da primeira pessoa do plural, isto é, são “aparentes” na oração, e por expressão nula os casos em que o pronome é omitido e indicado por meio das desinências *-mos* ou \emptyset .

Como hipótese, a autora entende que é a expressão plena (17) e não a expressão nula (18) que favorece a realização da forma *a gente*, uma vez que há a tendência desta forma acompanhar o verbo na terceira pessoa do singular, conforme visto em (19).

Ainda, Vitório (2017, p. 128) adverte que “essa forma verbal, em determinados contextos, promove ambiguidade de referente, por ser utilizada junto a uma vasta gama de pronomes pessoais”, portanto a expressão plena, em que o pronome fica marcado na oração, faz com que não ocorra confusão a respeito do referente.

(17) quando *nós* não *tínhamos* o armazém *a gente* tinha um descanso/ mais, mas/ depois que *nós* *abrimos* o armazém é vinte e quatro horas no ar. (L36L4892)

(18) *a gente* sai de casa \emptyset faz uma oração \emptyset pede muita proteção porque o índice de violência ele aumentou bastante *a gente* pode perceber no nosso dia a dia. (L30L4065)

(19) *a gente tem* um armazém de construção e tem um ano que a gente abriu/ mais, mas/ assim novo no comércio porque a gente não mexia nada. (L36L4781)

Fonte: Vitório (2017, p. 128)

Os resultados obtidos confirmam a hipótese de Vitório (2017), uma vez que a combinação *a gente* + expressão plena apresenta um percentual de 88% e um peso relativo de .79, enquanto que a combinação *a gente* + expressão nula apresenta um percentual mais baixo, de 68% e um peso relativo de .41. Tais resultados convergem com a literatura, como em Vitório (2015), que apontam uma maior frequência da combinação *a gente* + expressão plena, na medida em que a ausência de marca flexional do sujeito no verbo não serve para identificar a pessoa do discurso.

Passando para as variáveis extralinguísticas, a primeira selecionada foi o fator escolaridade. Vitório (2017) lança luz sobre a influência desta variável na preferência por uma variante, na medida em que, com base em Votre (2003), tem-se o entendimento de que “pessoas mais escolarizadas tendem a usar mais as formas prescritas nos manuais normativos”, isto é, quanto maior o nível de escolaridade do falante, maior será a preferência pelo uso das variantes consagradas nos materiais didáticos. Partindo deste princípio, a autora estabelece como hipótese a relação: quanto maior escolaridade, menor o uso da variante *a gente*.

Os valores para este fator também corroboram a hipótese da autora, apresentando percentuais de 93%, 86% e 78%, para os níveis de escolaridade fundamental, médio e superior, respectivamente. Em relação aos pesos relativos, tem-se como valores .71, para os falantes com ensino fundamental, .52, para os falantes do ensino médio, e .38, para os falantes do ensino superior, o que mostra um ponto de favorecimento por parte dos menos escolarizados, uma neutralidade nos falantes de nível médio, e um desfavorecimento entre aqueles que possuem um nível de escolaridade mais elevado.

A segunda variável extralinguística e a última considerada estatisticamente significativa na análise da variação entre *nós* e *a gente*, na função de sujeito, na fala maceioense, diz respeito à faixa etária, controlada, nas palavras da autora, para verificar se, na comunidade de fala estudada, tem-se um processo de variação estável ou de mudança em curso (LABOV, 1994). Vitório (2017) assume como pressuposto que a variante *a gente* é realizada mais frequentemente entre os falantes mais jovens, com estes valores percentuais diminuindo em direção aos grupos etários mais velhos.

Os valores encontrados nas faixas etárias F1 (15-29), F2 (30-44) e F3 (acima de 44 anos), controladas pela autora, mostram percentuais de 97% e 91%, respectivamente, para os

primeiros dois grupos, o que confirma um uso praticamente categórico/bastante elevado nestes dois grupos etários. Em relação ao terceiro grupo etário, dos mais velhos, tem-se um percentual de 69% para a realização de *a gente*, menor, mas que ainda apresenta uma preferência deste grupo pela realização da variante *a gente* na função de sujeito. Já nos pesos relativos, F1 apresenta o valor de .83, enquanto que F2 nos mostra um valor de .57, ambos favorecendo a realização de *a gente*. No entanto, o grupo etário F3, o mais velho da amostra, apresenta um valor de .22 em seu peso relativo, o que desfavorece o uso de *a gente*, sinalizando uma mudança em curso.

Ainda, a autora cruza os fatores *Faixa Etária e Escolaridade* e verifica que em todas as faixas etárias o aumento do nível de escolaridade faz com que o percentual de realização de *a gente* diminua, evidenciando o quão significativo este condicionador é na escolha entre a variante *a gente* e a variante *nós*. O contraste entre os valores de 100% e 99% encontrados para os falantes com ensino fundamental de F1 e F2, respectivamente, com os 64% do grupo F3, com falantes que possuem ensino superior, sinaliza uma curva de mudança em progresso (LABOV, 1994).

3 METODOLOGIA

Inserido na comunidade de fala de Porto Alegre, este estudo tem como objetivos principais (a) investigar o desenvolvimento do processo de variação entre as formas pronominais de primeira pessoa do plural, *nós* e *a gente*, na capital dos gaúchos, em um intervalo de 25 anos, com base na comparação de dados de duas amostras, uma coletada nos anos 1990, a do VARSUL, outra coletada entre 2015-2019, a do LínguaPOA, portanto, mais recente do que a utilizada em outros trabalhos, como o de Zilles (2005, 2007); e (b) testar a hipótese de que o pronome *nós* é mais realizado em contextos [+monitorados]: busca-se investigar se o fator *Estilo*, de fato, se correlaciona com a realização das duas formas pronominais na função de sujeito.

Para tal, esta pesquisa se utiliza de análise de regra variável em tempo real, estudo de tendência, comparando as duas amostras citadas acima: com dados de entrevistas sociolinguísticas da amostra do LínguaPOA (2015-2019), vinculadas ao projeto de pesquisa “Variação linguística e significados sociais no português falado em Porto Alegre (RS)”, sob a orientação da Profa. Dra. Elisa Battisti (UFRGS), do qual faço parte desde 2019; e em entrevistas sociolinguísticas da amostra do projeto VARSUL, organizado na UFRGS pela Profa. Dra. Leda Bisol, coletadas na década de 1990.

Neste capítulo, tem-se por objetivo apresentar os procedimentos metodológicos utilizados na pesquisa, assim como discorrer brevemente acerca dos acervos que compõem a análise, bem como a respeito da Porto Alegre de ontem em comparação com a Porto Alegre de hoje, neste intervalo de 25 anos.

3.1 A comunidade de fala de Porto Alegre

De acordo com dados do IBGE¹⁷, muito embora o estatuto de cidade só tenha sido concedido a Porto Alegre mais tardiamente, em 1821, por Dom Pedro I, a capital dos gaúchos teve seu processo de povoamento iniciado em 1752, em virtude do Tratado de Madri, que incentivou a imigração de 60 casais açorianos à região. Ainda, sua fundação data de março de 1772, com a criação da Freguesia de São Francisco do Porto dos Casais. No século XX, a cidade passou por um acelerado processo de desenvolvimento que a alçou e a mantém até hoje entre as principais capitais do país.

No último censo (2010), a população de Porto Alegre era de 1.409.351 pessoas, com uma densidade demográfica de 2.837,53 hab/km². Estima-se que, neste ano, em 2021, este

¹⁷ <https://cidades.ibge.gov.br/brasil/rs/porto-alegre/> Acesso em: 03/10/2021

número aumentará para 1.492.530 pessoas. Em relação aos aspectos étnicos, no censo de 2010, 79,23% dos porto-alegrenses autodeclararam-se brancos e 20,24% negros. Além disso, segundo o ObservaPOA¹⁸, Porto Alegre é a capital brasileira que apresenta o maior percentual de população idosa, com 15,04%, seguido de Rio de Janeiro (14,89%) e Belo Horizonte (12,61%).

Também cabe mencionar que este percentual de 15,04% representa o triplo da capital com menor percentual de população idosa, Palmas, com 4,37%, e se aproxima muito do triplo de capitais como Macapá (5,15%) e Boa Vista (5,18%), o que denota uma mudança significativa na pirâmide etária da cidade.

Segundo levantamento da FGV Social¹⁹, Porto Alegre figura na segunda posição entre as capitais mais ricas do país, com uma renda média mensal por habitante de R\$ 3.725,15. De acordo com Neri da FGV, entre os motivos para Porto Alegre figurar nesta posição estão “as aposentadorias do setor público, o alto capital humano e social e o fato de a capital gaúcha ter mais empregos formais do que a média nacional”. Conforme ilustra o Atlas do Desenvolvimento Humano (Figura 2), com base em dados do IBGE, é possível notar um crescimento em todos os índices relacionados ao IDHM (Índice de Desenvolvimento Humano) da capital dos gaúchos, em relação à média brasileira e a si próprio, desde o censo de 1991 até a última amostragem, em 2010, o que atesta sua posição entre uma das cidades mais desenvolvidas do território nacional.

Figura 2 – Índices de IDHM de Porto Alegre

Territorialidades	IDHM	IDHM	IDHM	IDHM	IDHM	IDHM						
	Censo	Censo	Censo	Renda	Renda	Renda	Longevidade	Longevidade	Longevidade	Educação	Educação	Educação
	1991	2000	2010	Censo	Censo	Censo	Censo	Censo	Censo	Censo	Censo	Censo
Brasil	0,493	0,612	0,727	0,647	0,692	0,739	0,662	0,727	0,816	0,279	0,456	0,637
Porto Alegre (RS)	0,660	0,744	0,805	0,779	0,830	0,867	0,748	0,811	0,857	0,494	0,612	0,702

Fonte: Atlas Brasil²⁰

Considerando-se o perfil socioeconômico de Porto Alegre, supõe-se a ocorrência de mudanças em todos os níveis dentro de uma comunidade de fala que vem se desenvolvendo década após década. A forma como as pessoas se comunicam e usam a linguagem pode sofrer os efeitos dessas mudanças macrosociais. É o que se espera constatar na análise da variação

¹⁸ <https://gauchazh.clicrbs.com.br/porto-alegre/noticia/2016/12/porto-alegre-e-a-capital-com-a-maior-proporcao-de-idosos-na-populacao-8653774.html> Acesso em: 03/10/2021

¹⁹ <https://gauchazh.clicrbs.com.br/colunistas/marta-sfredo/noticia/2020/08/onde-estao-os-ricos-do-brasil-e-do-rs-nao-exatamente-onde-se-imagina-ckdpxnlq2001m01478hk4f0as.html> Acesso em: 03/10/21

²⁰ <http://www.atlasbrasil.org.br/> Acesso em: 03/10/21

entre *nós* e *a gente* aqui realizada: uma mudança nas frequências de uso das variantes *nós* e *a gente* acompanhando as mudanças sociais. Com isso em mente, faz-se necessário investigar, nos limites deste trabalho, o quanto estes fatores socioeconômicos, em conjunto com os fatores linguísticos e sociais controlados na análise, influenciam na variação entre os pronomes *nós* e *a gente* desde a década de 1990 até a década de 2010.

3.2 As entrevistas sociolinguísticas do LínguaPOA e VARSUL

Formado na década de 1990 pela parceria entre quatro universidades brasileiras (UFRGS, PUCRS, UFSC e UFPR), o Projeto VARSUL (Variação Linguística na Região Sul do Brasil) tem por objetivo geral a descrição do português falado e escrito de áreas socioculturalmente representativas do Sul do Brasil. Como propósitos específicos, o VARSUL busca oferecer (1) subsídios para a descrição do português falado e escrito no Brasil; (2) condições para teste e desenvolvimento de teorias linguísticas; (3) condições para formação de novos pesquisadores; e (4) subsídios para programas educacionais, promovendo o conhecimento e o respeito às variedades linguísticas. Conforme consta no *site* do projeto²¹, o acervo base é constituído por 288 entrevistas, distribuídas igualmente entre os três estados, sendo 24 por município: Rio Grande do Sul (Porto Alegre, Flores da Cunha, Panambi e São Borja); Santa Catarina (Florianópolis, Blumenau, Lages e Chapecó) e Paraná (Curitiba, Londrina, Pato Branco e Irati).

Em relação à estratificação do banco de dados, foram levados em consideração o sexo/gênero sob uma perspectiva binária (masculino e feminino), escolaridade (nível fundamental I - de 1 a 4 anos de escolaridade, nível fundamental II - de 5 a 8 anos de escolaridade e nível médio - de 9 a 11 anos de escolaridade) e idade (de 25 até 50 anos e acima de 50). Ainda, os informantes precisavam a) falar apenas português (exigência para os entrevistados nas capitais, mas não nas áreas bilíngues); b) ter morado na cidade pelo menos 2/3 de sua vida; c) não ter morado fora da região por mais de um ano no período de aquisição da língua nativa.

Vinculado ao projeto “Variação fonético-fonológica e classe social na comunidade de fala de Porto Alegre”, sob a orientação da Profa. Dra. Elisa Battisti, e com dados coletados por pesquisadores do Instituto de Letras da UFRGS no período de 2015-2019, o LínguaPOA é um acervo que reúne dados de fala do português brasileiro falado em Porto Alegre (POA), capital

²¹ <http://www.varsul.org.br/> Acesso em: 03/10/21

do estado do Rio Grande do Sul. Como informado no site do LínguaPOA²², a amostra estratifica-se por zonas da cidade (Central, Norte, Sul e Leste), sendo 2 bairros por zona, um de classe alta e um de classe baixa (considerando-se a renda média domiciliar mensal em salários mínimos das famílias do bairro, cf. ObservaPOA www.observapoa.com.br), 3 grupos etários (20-39 anos, 40-59 anos, 60+ anos), 3 níveis de escolaridade (Fundamental, Médio e Superior) e uma perspectiva de sexo/gênero binária (masculino/feminino). O projeto tem os seguintes objetivos, conforme informado no site do LínguaPOA (<https://www.ufrgs.br/linguapoa/>):

- a) investigar a configuração da comunidade de fala de Porto Alegre em termos de normas e características linguísticas partilhadas nas diferentes áreas do espaço geográfico e em seus estratos sociais;
- b) esclarecer a estruturação de Porto Alegre em classes sociais, relacionando distinção social à padronização e distinção linguística;
- c) verificar as variáveis linguísticas e extralinguísticas condicionadoras da aplicação de processos fonético-fonológicos variáveis no português brasileiro falado em Porto Alegre.

O acervo do LínguaPOA (2015-2019) contém 103 entrevistas sociolinguísticas. Conforme Freitag *et al.* (2021, p. 9), “as 103 entrevistas atendem a todos os critérios de estratificação nos níveis médio e superior de escolaridade, mas não no fundamental”, nível esse que, diferentemente do previsto, foi contemplado com apenas 7 entrevistas das 36 planejadas.

3.3 Análise de regra variável

Esta seção trata da análise de regra variável, no que diz respeito às amostras utilizadas (3.3.1), os procedimentos metodológicos (3.3.2) e as variáveis envolvidas (3.3.3).

3.3.1 As amostras LínguaPOA e VARSUL

As amostras aqui selecionadas coincidem apenas em parte com Zilles (2007), que utilizou dados da amostra do Projeto VARSUL nas etapas do seu trabalho acerca da variável-resposta investigada nesta pesquisa. Foram selecionadas para o presente estudo 12 entrevistas sociolinguísticas do LínguaPOA (2015-2019), com os informantes divididos em três faixas etárias (20-39, 40-59, 60 anos ou mais) e gênero binário (masculino e feminino), com o objetivo de comparar com outras 12 entrevistas sociolinguísticas do Projeto VARSUL, seguindo a mesma estratificação social, como pode se observar nos quadros 1 e 2.

²² <https://www.ufrgs.br/linguapoa/> Acesso em: 03/10/21

Quadro 1 – Os 12 informantes do LínguaPOA

Informante	Gênero	Faixa Etária
06	Feminino	20-39
60	Feminino	20-39
93	Masculino	20-39
111	Masculino	20-39
30	Feminino	40-59
120	Feminino	40-59
63	Masculino	40-59
81	Masculino	40-59
90	Feminino	60+
126	Feminino	60+
15	Masculino	60+
69	Masculino	60+

Fonte: autoria própria

Quadro 2 – Os 12 informantes do VARSUL

Informante	Gênero	Faixa Etária
08	Feminino	20-39
12	Feminino	20-39
11	Masculino	20-39
15	Masculino	20-39
02	Feminino	40-59
05	Feminino	40-59
01	Masculino	40-59
03	Masculino	40-59
16	Feminino	60+
24	Feminino	60+
07	Masculino	60+
18	Masculino	60+

Fonte: autoria própria

3.3.2 Os procedimentos metodológicos

Para a análise de regra variável, os dados foram levantados mediante a oitiva das 24 entrevistas (12 do LínguaPOA e 12 do VARSUL) utilizadas e registrados em uma planilha do Google Drive, para que, posteriormente, pudessem ser codificados em função da variável-resposta (realização de *nós* ou *a gente* na função de sujeito) e das variáveis previsoras (linguísticas e extralinguísticas), mais profundamente apresentadas na próxima seção.

Durante a oitiva dessas entrevistas, foi considerado o tempo integral de cada uma delas, visto que a quantidade de dados da variável-resposta disponíveis em cada entrevista se mostrou muito baixo. Além disso, dados em que *nós* ou *a gente* foram realizados numa função que não a de sujeito não compõem a presente análise. Ainda, é preciso salientar que os dados referentes à variável previsora *Estilo* foram coletados de duas formas distintas: na amostra do LínguaPOA (2015-2019) mediante a oitiva e a partir da anotação em uma trilha denominada *Estilo*, no programa Elan²³, e na amostra do VARSUL apenas mediante oitiva, pois não foi possível ter acesso às transcrições no período de realização desta pesquisa.²⁴

Uma vez que a metodologia aplicada aqui é de natureza quantitativa, os dados codificados foram submetidos a uma análise estatística no programa computacional R²⁵, versão 3.6.0, na interface RStudio, em modelos de regressão logística, para que se pudesse verificar quais variáveis previsoras se correlacionam à variável-resposta. A análise da variável-resposta (variação entre *nós* e *a gente*) foi concebida em uma perspectiva binominal (realização de *nós* ou de *a gente* na função de sujeito), tendo como etapas (1) a realização do teste de qui-quadrado (de Pearson) das variáveis previsoras nas duas amostras; e (2) uma análise multivariada de regressão logística, modelo linear geral com efeitos mistos, utilizando a função *glmer* do R, na qual foram consideradas todas as variáveis previsoras com p-valor significativo (igual ou menor do que 0,05) encontrados na etapa 1 e Informante como variável aleatória.

²³ <https://archive.mpi.nl/tla/elan> Acesso em: 03/10/21

²⁴ A pesquisa que resultou no presente Trabalho de Conclusão de Curso, sobre a variação *nós~a gente*, foi realizada conforme os Planos de Trabalho de bolsista de Iniciação Científica executados por este autor sob orientação da Profa. Elisa Battisti ao longo de 2020 e 2021, durante a pandemia de Covid-19. A orientadora, pesquisadora do VARSUL, contava com os áudios digitalizados das 12 entrevistas usadas no presente estudo antes de a pandemia eclodir. Já a consulta às transcrições das entrevistas do VARSUL, que só poderia ocorrer presencialmente, no Instituto de Letras da UFRGS, não foi possível porque a Universidade funcionou apenas de forma remota durante todo o período. O acervo do LínguaPOA (áudios e transcrições), diferentemente, está todo digitalizado, por isso a utilização remota foi possível.

²⁵ <https://www.rstudio.com/> Acesso em: 03/10/21. “O R é uma linguagem de programação voltada à análise de dados, que pode ser utilizada para realizar computações estatísticas e gráficas, compilar e anotar corpora, produzir listas de frequências, entre diversas outras tarefas”. (FREITAG, 2014, p. 134)

3.3.3 As variáveis

3.3.3.1 Variável-Resposta

A variável-resposta desta pesquisa é a variação entre *nós* e *a gente* na função de sujeito (*a gente gosta de passear/a gente gostamos de passear ~ nós gostamos de passear/nós gosta de passear*) e o fator de aplicação nos modelos estatísticos é a realização *nós*²⁶. Como já mencionado, foram consideradas apenas as realizações das duas variantes na função de sujeito. Também é importante destacar que realizações ditongadas, como *nóis*, estão entre os dados analisados. Ainda, em relação à variante *a gente*, as realizações *ahente*, *a'ente* e *'nte*, discutidas em Zilles (2007), e que ilustram a redução fonética de *a gente*, não foram consideradas/observadas nesta análise.

3.3.3.2 Variáveis previsoras

As variáveis previsoras controladas foram definidas com base em estudos anteriores acerca da variável-resposta (ZILLES, 2005, 2007; VITÓRIO, 2017) e da alternância estilística (BAUGH, 2001; ECKERT, 2001; SCHILLING-ESTES, 2007; VALLE e GÖRSKI, 2014), tendo como resultado a definição de três variáveis previsoras linguísticas e duas variáveis previsoras extralinguísticas (ou sociais).

3.3.3.2.1 Variáveis linguísticas

a) Marca Morfêmica

Leva-se em consideração a combinação *a gente/nós* + verbo, tendo como opções possíveis *a gente* + morfema zero (*a gente vai*) e *a gente* + morfema *-mos* (*a gente vamos*), assim como *nós* + morfema zero (*nós vai*) e *nós* + morfema *-mos* (*nós vamos*). Tem-se a hipótese de que o pronome *a gente* costuma ser realizado com maior frequência na combinação *a gente* + morfema zero, com base em estudos anteriores, como o de Zilles (2007).

²⁶ O *script* usado neste trabalho para efetuar a regressão logística na análise binomial (realiza vs. não realiza uma das formas de interesse) com o programa R selecionou como “fator de aplicação” o segundo dos fatores da variável-resposta (i.e., o segundo dos níveis da coluna referente à variável-resposta na planilha de dados) considerando-se seu ordenamento alfabético. Vale dizer, os fatores da variável-resposta são ‘a.gente’ e ‘nos’ na planilha de dados (ver figuras 6 e 7 adiante), por isso o R escolheu ‘nos’ como a realização cuja probabilidade de ocorrência foi estimada face ao conjunto de variáveis explanatórias (variáveis-previsoras) controladas.

b) *Paralelismo Formal*

Trabalha-se aqui com a ideia de que, iniciada uma sequência discursiva optando-se por uma das duas variantes (*nós* ou *a gente*), espera-se que se continue utilizando essa variante até o fim da sequência. Temos como fatores nessa variável: realização isolada, primeira da série, antecedida por *nós* e antecedida por *a gente*, conforme já ilustrado nos exemplos (13), (14), (15) e (16), respectivamente.

c) *Estilo*

Em relação à variável previsora *Estilo*, leva-se em consideração aqui que a alternância estilística na fala se dá em um *continuum* que vai da fala [-monitorada] à [+monitorada] (LABOV, 2008). Para viabilizar a análise em função do número relativamente reduzido de dados, foram levados em consideração os nós contextuais da Árvore de Decisão (LABOV, 2001), mas lida-se com a variável de maneira binária²⁷ na codificação dos dados, conforme os trechos em que ocorre cada realização de *nós* ou *a gente* na função de sujeito: opõem-se Estilo Monitorado (*Resposta, Língua, Soapbox, Residual*) a Estilo Espontâneo (*Narrativa, Grupo, Infância, Tangente*).

3.3.3.2.2 Variáveis extralinguísticas (ou sociais)

a) *Gênero*

A variável *Gênero* foi controlada sob uma perspectiva binária (Masculino e Feminino), categorização adotada tanto na amostra do VARSUL quanto na amostra do LínguaPOA. Controla-se Gênero com o objetivo de verificar dois importantes pressupostos labovianos (LABOV, 2008), nos quais (1) as mulheres costumam liderar a inserção de variantes inovadoras na comunidade de fala; e (2) as mulheres costumam rejeitar o uso dessas variantes na medida em que passam a receber um significado social negativo. No que se refere a *nós* e *a gente*, a hipótese testada neste trabalho é a de que se trata de uma mudança liderada pelas mulheres e que ambos os gêneros apresentam uma taxa maior de realização de *nós* nos contextos monitorados.

²⁷ Aqui, ao definir a forma como lidou-se com a variável previsora por “binária”, faz-se referência a divisão dos oito nós contextuais da Árvore de Decisão (Labov, 2001) em dois eixos estilísticos: monitorado ou espontâneo. Para ilustrar: quando encontrada uma realização de *nós* ou *a gente*, por exemplo, em um contexto de nó resposta, essa realização era codificada como monitorada. Em contrapartida, se a realização de uma das variantes era encontrada em um contexto de nó narrativa ou outro qualquer pertencente ao eixo casual da Árvore de Decisão Laboviana, tal ocorrência era registrada como espontânea.

b) Faixa Etária

Controla-se a faixa etária dos informantes das duas amostras (LínguaPOA e VARSUL), dividindo-se os em três grupos etários: 20-39, 40-59, 60 ou mais anos, com o objetivo de verificar como se manifesta a variação entre as formas pronominais *nós* e *a gente*, na função de sujeito, na medida em que se passa de um grupo etário a outro. A hipótese deste trabalho é a de que o uso de *a gente* na função de sujeito se trata de uma mudança em progresso no português de Porto Alegre, com aumento na proporção de uso dessa variante à medida que a idade dos informantes diminui.

3.3.4 A codificação dos dados

No que diz respeito à codificação dos dados de *nós* e *a gente*, tal tarefa se divide em duas etapas: análise das entrevistas sociolinguísticas do LínguaPOA via oitiva e com o suporte das transcrições na ferramenta Elan (no VARSUL apenas oitiva) e anotação das ocorrências de *nós* e *a gente* em uma planilha de dados do Google Drive, conforme figuras (3), (4) e (5).

Figura 3 – Planilha de dados *nós* e *a gente* (LínguaPOA)

NOS.A.GENTE	INFORMANTE	MARCA.MORF	PARALEL.FORM	ESTILO		GENERO	FAIXA.ETARIA
a.gente	inf81	zero	isolada	Monitorado	00:22	Masculino	40-59
a.gente	inf81	zero	isolada	Monitorado	01:28	Masculino	40-59
a.gente	inf81	zero	primeira	Monitorado	02:05	Masculino	40-59
a.gente	inf81	zero	antecedida.a.gei	Monitorado	02:20	Masculino	40-59
a.gente	inf81	zero	antecedida.a.gei	Monitorado	02:23	Masculino	40-59
a.gente	inf81	zero	isolada	Monitorado	02:33	Masculino	40-59
a.gente	inf81	zero	primeira	Espontâneo	03:08	Masculino	40-59
a.gente	inf81	zero	antecedida.a.gei	Espontâneo	03:16	Masculino	40-59
a.gente	inf81	zero	antecedida.a.gei	Espontâneo	03:28	Masculino	40-59
a.gente	inf81	zero	antecedida.a.gei	Espontâneo	03:41	Masculino	40-59
a.gente	inf81	zero	antecedida.a.gei	Espontâneo	03:47	Masculino	40-59
a.gente	inf81	zero	primeira	Monitorado	07:58	Masculino	40-59
a.gente	inf81	zero	antecedida.a.gei	Monitorado	08:04	Masculino	40-59
a.gente	inf81	zero	primeira	Monitorado	09:07	Masculino	40-59
a.gente	inf81	zero	antecedida.a.gei	Espontâneo	09:17	Masculino	40-59
nos	inf81	mos	antecedida.a.gei	Espontâneo	09:20	Masculino	40-59

Fonte: autoria própria

Figura 4 – Planilha de dados *nós* e *a gente* (VARSUL)

NOS.A.GE	INFORMA	MARCA.M	PARALEL.FORM	ESTILO		GENERO	FAIXA.ETARIA
nos	inf1	mos	primeira	Monitorado	00:23	Masculino	40-59
nos	inf1	mos	antecedida.nos	Monitorado	00:32	Masculino	40-59
nos	inf1	mos	primeira	Monitorado	02:02	Masculino	40-59
nos	inf1	mos	antecedida.nos	Espontâneo	02:11	Masculino	40-59
nos	inf1	mos	primeira	Espontâneo	02:40	Masculino	40-59
nos	inf1	mos	antecedida.nos	Espontâneo	02:43	Masculino	40-59
nos	inf1	mos	antecedida.nos	Espontâneo	02:55	Masculino	40-59
nos	inf1	mos	antecedida.nos	Espontâneo	03:06	Masculino	40-59
nos	inf1	mos	antecedida.nos	Espontâneo	03:13	Masculino	40-59
nos	inf1	mos	isolada	Monitorado	03:26	Masculino	40-59
nos	inf1	mos	isolada	Monitorado	04:02	Masculino	40-59
nos	inf1	mos	isolada	Monitorado	05:58	Masculino	40-59
nos	inf1	mos	isolada	Monitorado	06:33	Masculino	40-59
a.gente	inf1	zero	isolada	Monitorado	07:21	Masculino	40-59
nos	inf1	zero	isolada	Monitorado	07:29	Masculino	40-59
nos	inf1	zero	primeira	Espontâneo	08:43	Masculino	40-59

Fonte: autoria própria

Figura 5 – Janela de transcrição em ELAN de uma entrevista do LínguaPOA

The screenshot shows the ELAN transcription interface. At the top, there are two audio waveforms. Below them is a time axis from 12.000 to 00:00:24.000. The transcription grid below the axis is divided into several segments:

- D1 [100]**: A red segment containing the text "e vocês mo".
- S [934]**: A purple segment containing the text "a família é | a família próxima assi | mais nov | e isso pai e mã | ã não a gente mora sou c".
- Dados conte [17]**: A green segment.
- Roteiro [13]**: A yellow segment.
- D2 [28]**: A green segment.
- Estilo [118]**: A blue segment containing the text "esposta" and "Resposta".

Fonte: autoria própria

Nas Figuras 3 e 4, apresentam-se exemplos na planilha de dados da codificação de *nós* e *a gente* realizada na análise de 12 entrevistas sociolinguísticas das amostras do LínguaPOA e do VARSUL. Em cada coluna das planilhas, os dados são codificados conforme cada variável controlada: *Nós* e *A gente* (realização variável), *Informante* (entre os(as) 12 de cada amostra), *Marca Morfêmica*, (zero ou -mos), *Paralelismo Formal* (isolada, primeira, antecedido por a gente e antecedido por nós), *Estilo* (Monitorado e Espontâneo) e os fatores *Gênero* (masculino e feminino) e *Faixa Etária* (20-39, 40-59, 60+) sendo controlados.

Já na figura 5, temos uma janela de transcrição da ferramenta Elan²⁸, utilizada exclusivamente com a amostra do LínguaPOA, em que é possível observar uma realização de *a gente* na função de sujeito e também a trilha ESTILO com a codificação “resposta”, que, num contexto de alternância estilística binária, seria codificado como “Monitorado” nas planilhas (3) e (4), com base em Labov (2001).

²⁸ Explicação sobre os códigos das trilhas na ferramenta Elan: D1 (entrevistador(a)), S (informante), Dados Contextuais (anotação de pistas do canal (Labov, 2008), como risos, ou até mesmo menção a barulhos externos que estejam ocorrendo no ambiente), Roteiro (o assunto geral em discussão em cada parte da entrevista), D2 (trilha dedicada a outro falante que esteja interagindo na situação de entrevista, o que nem sempre ocorre), Estilo (espaço dedicado a anotação dos nós contextuais da Árvore de Decisão (Labov, 2001)).

4 APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

Neste capítulo, tem-se por objetivo apresentar e discutir os resultados encontrados para as amostras do LínguaPOA e do VARSUL. Para tal tarefa, o capítulo se dividirá em ordem cronológica, partindo-se dos resultados da análise dos dados da amostra mais antiga (VARSUL) e em seguida apresentando-se os resultados encontrados na análise da amostra mais recente (LínguaPOA).

4.1 Análise de regra variável (Resultados)

Nesta seção, serão explorados os resultados estatísticos encontrados via *software* R (4.1.1 e 4.1.2) para a amostra do VARSUL (1990) e a amostra do LínguaPOA (2015-2019), buscando apresentar e discutir a variação entre as formas pronominais *nós* e *a gente*, na função de sujeito, no português de Porto Alegre em dois momentos para, assim, cumprir o primeiro de nossos objetivos, como informado na introdução: verificar se houve alteração nas proporções de uso de *nós* e *a gente* na comunidade em cerca de 25 anos. A esse respeito, é importante reforçar: nos modelos de regressão logística cujos resultados apresentaremos nas próximas tabelas, a realização *nós* corresponde ao fator de aplicação.

4.1.1 Resultados Estatísticos – VARSUL

Na amostra de entrevistas do VARSUL, foram levantadas e codificadas 593 realizações de *nós* ou *a gente* na função de sujeito, sendo 35% destes dados referentes à variante *nós* e os outros 65% referentes à variante *a gente*. Tais resultados se aproximam dos apresentados em Zilles (2007), que encontrou 72% de realização de *a gente*, na função de sujeito, nos dados do VARSUL (1990), em sua análise em tempo real/estudo de tendência entre esta amostra e a do NURC (1970).

Na tabela 3, estão os resultados para a amostra do VARSUL. O modelo inclui apenas as variáveis que, no teste de qui-quadrado, apresentaram p-valores significativos: *Paralelismo Formal, Estilo, Gênero, Faixa Etária*.

No que diz respeito às variáveis controladas nesta pesquisa, podemos perceber que dois dos grupos de fatores apresentam correlação com a realização da variante *nós*: *Paralelismo Formal e Gênero*.

Em relação ao *Paralelismo Formal*, os achados convergem com o apresentado em Vitorio (2017), na medida em que *antecedido por a gente*, valor de referência, tem menor proporção de realização do que *nós*, e as estimativas de todos os demais fatores têm p-valor significativo e são positivas (favorecem *nós*).

A outra variável no modelo da Tabela 3 que exibe estimativa com p-valor significativo é o *Gênero*. O fator Masculino exibe a maior proporção de realização de *nós* e a estimativa confirma seu papel favorecedor. De certa forma, os resultados convergem com os de Zilles (2007), em que as mulheres apresentam uma proporção maior de realização da variante *a gente* e menor de *nós* na comparação com os homens.

A respeito da *Marca Morfêmica*, essa variável não foi incluída no modelo porque, apesar de apresentar p-valor significativo, no teste de qui-quadrado, foram encontrados os valores de 0% para combinação *a gente* + *-mos* e 100% para a combinação *nós* + *-mos*, o que compromete as estimativas da variável *Marca Morfêmica*. Tais resultados estão de acordo com os encontrados em Zilles (2007) em que a combinação entre a variante *a gente* e o morfema zero é praticamente categórica.

Tabela 3 - Estimativas dos parâmetros do modelo (de regressão logística, modelo linear geral com efeitos mistos) da realização de *nós* na função de sujeito.

N = 593

Intercepto = -2.7221

Variável	Apl./Ocorrências	Estimativa	Erro padrão	z	P
Paralelismo formal					
Antec.a.gente (valor de ref.)	15/178 (8,4%)				
Antec.nos	69/91 (75,8%)	2.5316	0.3992	6.342	2.27e-10 ***
Isolada	64/177 (36,1%)	1.2912	0.3453	3.739	0.000185 ***
Primeira	60/147 (40,8%)	1.5698	0.3443	4.560	5.11e-06 ***
Estilo					
Espontâneo (valor de ref.)	86/203 (42,3%)				
Monitorado	122/390 (31,2%)	-0.2509	0.2407	-1.042	0.297374
Gênero					
Feminino (valor de ref.)	62/349 (17,7%)				
Masculino	146/244 (59,8%)	1.8076	0.5092	3.550	0.000385 ***
Faixa Etária					
20-39 (valor de ref.)	54/144 (37,5%)				
40-59	116/239 (48,5%)	0.7312	0.6141	1.191	0.233794
60+	38/210 (18%)	-0.5317	0.6320	-0.841	0.400210

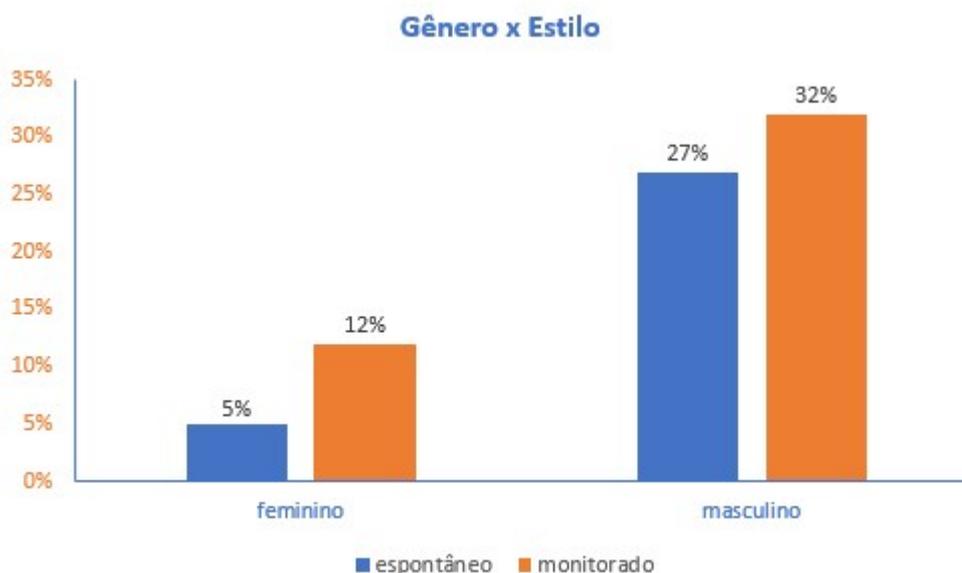
Modelo 1. (NOS.A.GENTE ~ PARALEL.FORM + ESTILO + GENERO + FAIXA.ETARIA + (1|INFORMANTE))

Fonte: autoria própria

A variável *Estilo*, incluída no modelo, não apresentou estimativas com p-valor significativo, ou seja, *Estilo* não se correlaciona à realização de *nós* na amostra VARSUL, como esperávamos. Face a esse resultado, fez-se uma primeira exploração das razões para a ausência

de correlação. Promoveu-se um cruzamento entre esse fator e *Gênero* para a variável-resposta *nós*, buscando verificar efeito conjunto dessas variáveis. Os resultados de tal cruzamento, apresentados no gráfico 4, expõem um panorama mais claro a respeito dessa relação. Podemos perceber que os homens são mais conservadores e preferem a variante *nós* tanto no estilo espontâneo (mulheres 5% vs. homens 27%) quanto no monitorado (mulheres 12% vs. homens 32%). Já as mulheres usam menos *nós* porque, inversamente, empregaram mais *a gente*. Além disso, ambos os gêneros aumentam a sua realização de *nós* nos contextos monitorados (mulheres 5% > 12% e homens 27% > 32%). Isso mostra por que o fator *Estilo* não se correlaciona ao uso de *nós*: independentemente do Estilo, os homens são mais conservadores e apresentam uma proporção maior de *nós*.

Gráfico 4 – Amostra VARSUL – Realização de *nós* por *Gênero* e *Estilo*



Fonte: autoria própria

Para ilustrar os resultados do cruzamento entre os fatores *Gênero* e *Estilo*, é válido trazer para a discussão a simbólica entrevista do informante 1 da Amostra do Varsul. Desde o início da entrevista, o informante 1, do sexo masculino e da faixa etária dos 40 aos 59 anos, deu sinais de ser uma pessoa falante e que suas respostas não seriam restritas ao preenchimento de um roteiro pré-estabelecido. A primeira impressão de que a conversa seria mais leve, algo próximo da fala vernacular tão desejada pelos sociolinguistas se confirma, mas chama muito a atenção a preferência do informante pela variante *nós*. Em toda a entrevista, embora emergjam os trechos narrativos longos e muito divertidos, o informante 1, de suas 68 realizações de *nós* ou *a gente*, utiliza apenas três vezes a variante gramaticalizada, o que representa apenas 4,4%

de *a gente* em toda a entrevista, e dessas realizações todas aparecem em contextos [+monitorados].

Podemos observar, por exemplo, parte de um trecho narrativo desta entrevista em que o informante opta pela utilização da variante conservadora:

Aí o rapaz que tava na rodoviária nos esperando chega pra nós // olha tem uma zebra pra vocês // eu digo aí pera aí outra qual é // os hotel tão tudo lotado // e aí eu disse mas e agora meu filho? // não, não se preocupa, **nós** vamo dar um jeito, **nós** vamo arrumar essas pensão aí, dormitório, dos caminhoneiro. (VARSUL, INF5)

Já no caso da entrevista da informante 5, é possível observar uma preferência pelo uso da variante *a gente* nos contextos [-monitorados]. Em um momento da entrevista, em que a temática introduzida é a da religião. A informante até inicia o trecho seguindo o curso da pergunta, mas logo realiza uma tangente, fazendo um relato sobre seu pai, que a família pensa estar doente, e se utiliza da variante gramaticalizada para isso:

Ele já veio de Santa Catarina com aquela dor aqui, no nervo ciático parece que foi, e aí ele começou, mas ele não diz. Meu pai é o tipo da pessoa que fica doente, mas tu não sabe, ele só começa. Agora por exemplo ele tá magro, ele tá emagrecendo, mas eu acho, eu creio que seja duma tosse que ele tá, ele fuma né e **a gente** vê que aquilo não se desprende aquele catarro nos pulmão. (VARSUL, INF1)

4.1.2 Resultados Estatísticos – LínguaPOA

Avançando mais de 20 anos no tempo, temos os resultados da análise da amostra do LínguaPOA. Nessa amostra, foram levantadas e codificadas 873 realizações de *nós* ou *a gente* na função de sujeito, sendo 13% destas referentes à variante *nós* e 87% referentes à variante *a gente*. Tais achados convergem com o esperado por Zilles (2007) na medida que (a) *a gente* é mais frequente do que *nós* e (b) confirma-se a tendência constatada pela autora de que o uso de *a gente* progrediria na comunidade. Houve uma diminuição significativa na realização da variante *nós* (35% > 13%) da amostra da década de 1990 em comparação com a amostra do LínguaPOA, mais recente.

No que diz respeito às variáveis previsoras apontadas como correlacionadas à realização de *nós* pelo R (tabela 4), apenas o *Paralelismo Formal* tem esse papel. Na comparação com os resultados do VARSUL, o *Paralelismo Formal* continua se manifestando da mesma forma, com os fatores *primeira da série*, *isolada* e *antecedido por nós* favorecendo o uso da variante *nós*.

Já os fatores sociais *Gênero* e *Faixa Etária* não se correlacionam ao uso de *nós*, o que é peculiar a processos de mudança em progresso já bastante avançados.

Sobre o fator *Estilo*, tem-se a hipótese de que o percentual elevado de realização de *a gente* (87%) possa estar reduzindo a influência estilística no ambiente de entrevista, o que também aponta uma inserção cada vez maior da variante *a gente* em contextos que, *a priori*, geram um grau de monitoramento maior por parte dos indivíduos.

Tabela 4 - Estimativas dos parâmetros do modelo (de regressão logística, modelo linear geral com efeitos mistos) da realização de nós na função de sujeito.

N = 873

Intercepto = -5.1174

Variável	Apl./Ocorrências	Estimativa	Erro padrão	z	P
Paralelismo Formal					
Antec.a.gente (valor de ref.)	17/404 (4,2%)				
Antec.nos	26/52 (50%)	2.1333	0.3986	5.353	8.67e-08 ***
Isolada	34/186 (18,2%)	1.2381	0.3432	3.607	0.000310 ***
Primeira	33/231 (14,2%)	1.1906	0.3315	3.592	0.000329 ***
Gênero					
Feminino (valor de ref.)	34/416 (8,1%)				
Masculino	76/457 (16,6%)	0.1998	0.9098	0.220	0.826220
Faixa etária					
20-39 (valor de ref.)	20/376 (5,3%)				
40-59	28/267 (10,4%)	1.3431	1.1733	1.145	0.252331
60+	62/230 (26,9%)	2.2265	1.1759	1.893	0.058310

Modelo 1. (NOS.A.GENTE ~ + PARALEL.FORM + GENERO + FAIXA.ETARIA + (1|INFORMANTE))

Fonte: autoria própria

4.2 A variação nós~a gente em tempo real

Em posse dos resultados apresentados acima para as amostras do VARSUL e LínguaPOA, faz-se necessário discutir de que modo esses indivíduos variam no decorrer de suas vidas. De acordo com Labov (1994, p. 83), “a interpretação de dados em tempo real, extraídos de estudos de painel ou tendência, requerem um modelo subjacente acerca de como os indivíduos variam ou não variam através de suas vidas”, isto é, quais são os padrões verificáveis nestes indivíduos e na comunidade. Para esclarecer a questão, o autor apresenta quatro padrões possíveis de variação nos indivíduos/comunidade no decorrer do tempo: estabilidade, *age-grading*, mudança geracional, mudança comunal. A respeito do primeiro padrão (estabilidade), ocorre quando o comportamento dos indivíduos se mantém no decorrer de suas vidas e não há mudança na comunidade. O segundo padrão (*age grading*) é verificado em situações em que os indivíduos modificam seu comportamento linguístico no decorrer de

suas vidas, mas o mesmo não ocorre na comunidade. Já em relação ao terceiro e quarto padrões, Labov (1994) salienta que estes não são tão transparentes quanto os dois primeiros.

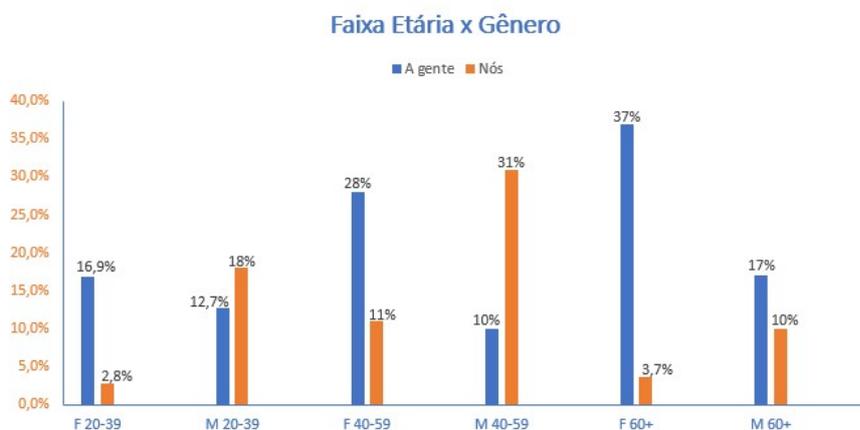
No que diz respeito a mudança geracional – mais recorrentes em mudanças fonológicas e morfológicas –, o autor aponta:

falantes individuais se inserem na comunidade com uma frequência de uso característica para determinada variável, o que se mantém no decorrer de suas vidas; mas o aumento regular nos valores adotados pelos demais indivíduos, frequentemente incrementados de geração para geração, acarretam em uma mudança linguística na comunidade. (LABOV, 1994, p. 84)

Referente ao quarto e último padrão, a mudança comunal, Labov (1994, p. 84) a define como um processo em que “todos os indivíduos da comunidade alteram suas frequências juntos, ou adquirem novas formas simultaneamente”, como no caso das comunidades bilíngues apresentadas no gráfico 2, que, num ritmo mais lento do que nas capitais, aumentam suas taxas de realização de *a gente* de maneira gradual.

Com isto posto, foi necessário realizar mais um cruzamento nas duas amostras, desta vez entre os fatores *Faixa Etária* e *Gênero*, com o intuito de esclarecer algumas questões que permaneceram pendentes após a apresentação dos demais resultados para, então, definir com qual tipo de mudança estamos lidando. Em relação a amostra do VARSUL, os resultados observados no cruzamento entre *Faixa Etária* e *Gênero* (gráfico 5) adicionam um novo aspecto ao cruzamento entre *Gênero* e *Estilo* (gráfico 4), visto que se verifica uma preferência por parte dos homens da variante *nós* nas faixas etárias mais jovens, mas não entre os mais velhos (17% de *a gente* vs. 10% de *nós*). Ainda, em comparação com as mulheres, os resultados mostram um movimento inverso. É possível observar um decréscimo no uso de *nós* da faixa etária mais velha para a mais jovem (60 anos aos 20-39) por parte do sexo feminino (3,7% > 2,8%), e um aumento deste número por parte dos homens (10% > 18%).

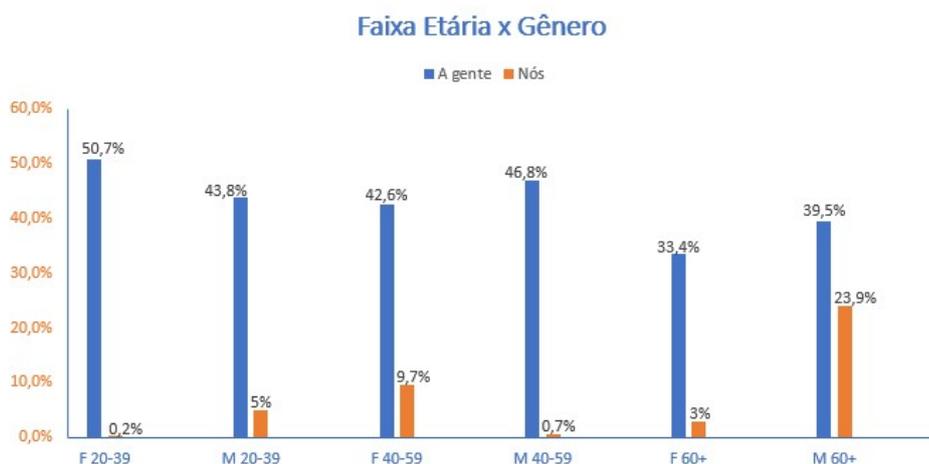
Gráfico 5 - Amostra VARSUL – Realização de *nós* e *a gente* por *Faixa Etária* e *Gênero*



Fonte: autoria própria

Já no cruzamento entre *Faixa Etária* e *Gênero* (gráfico 6), realizado na amostra do LínguaPOA, os resultados apontam para um processo de mudança bastante avançado tanto para os homens quanto para as mulheres. É possível perceber que a realização de *nós* diminui de 23,9% para 5%, no sexo masculino, da faixa etária mais velha (60+) para mais jovem (20-39), o que também ocorre com as mulheres, que apresentam um decréscimo de 3% para 0,2% de realização da variante *nós* nesta mesma direção. Tais resultados apresentam um cenário de mudança geracional com as mulheres uma geração à frente no processo, conforme previsto por Labov (1994).

Gráfico 6 - Amostra LínguaPOA – Realização de *nós* e *a gente* por *Faixa Etária* e *Gênero*



Fonte: autoria própria

5 CONCLUSÃO

Nesta pesquisa buscou-se identificar a variação entre as formas pronominais *nós* e *a gente*, na função de sujeito, e a influência dos estilos contextuais, na comunidade de fala de Porto Alegre, em um intervalo de 25 anos. Os resultados estatísticos encontrados nas duas amostras analisadas (VARFUL e LínguaPOA) nos mostram um cenário significativamente distinto para a variável em estudo dos anos 90 pra cá, com achados interessantes, mas que ainda necessitam ser afinados em etapas futuras de pesquisa com a inserção de condicionadores como Escolaridade (especialmente informantes de nível fundamental), bem como a realização de uma análise em tempo aparente.

Os resultados encontrados na amostra do LínguaPOA mostram que o uso de *a gente* está em estágio avançado na fala dos porto-alegrenses e que se trata de uma mudança em progresso, o que já sugeriam os resultados apresentados em Zilles (2007) em sua análise com a amostra do VARFUL (1990). Além disso, o fator *Paralelismo Formal* se mostra como significativo na variação entre *nós* e *a gente*, apontando uma preferência dos informantes por continuar utilizando a mesma variante com a qual iniciaram determinada sequência discursiva.

Especialmente em relação à questão estilística, a análise da amostra do LínguaPOA não comprova a correlação com a variável *Estilo*, que nem mesmo foi incluída no modelo de regressão logística porque, na análise de qui-quadrado²⁹, não se constatou diferença estatisticamente significativa entre os fatores [-monitorado] e [+monitorado]. Têm-se a hipótese de que, em uma amostra com um número maior de informantes/dados, em que se possa analisar esta variável previsora valendo-se dos oito nós contextuais da Árvore de Decisão (Labov, 2001), o que não foi feito neste estudo (dividimos os oito nós contextuais em Monitorado ou Espontâneo), a correlação entre *Estilo* e a *Variável-Resposta* possa ser diferente do apresentado no estudo aqui realizado. Salienta-se, também, a possibilidade de incorporar outras variáveis predictoras ao processo, como o fator Escolaridade, em etapas futuras de análise com a amostra do LínguaPOA, desta vez em tempo aparente.

Já nos resultados encontrados na amostra do VARFUL, além do fator *Paralelismo Formal*, o fator *Gênero* correlaciona-se à variação entre *nós* e *a gente*, com um percentual mais elevado de *nós* para os homens. A variável *Estilo* foi incluída no modelo de regressão logística porque o teste de qui-quadrado constatou diferença estatisticamente significativa entre os

²⁹ O p-valor alcançado no teste de qui-quadrado foi de 0,8547, as proporções de dados por fator foram: *a gente* [-monitorado] (191/220, ou 87%), *a gente* [+monitorado] (572/653, ou 88%), *nós* [-monitorado] (29/220, ou 13%), *nós* [+monitorado] (81/653, ou 12%).

fatores [-monitorado] e [+monitorado]. No entanto, a análise de regressão não comprova a correlação de *Estilo* com a realização de *nós* na amostra do VARSUL.

No que diz respeito à comparação desta pesquisa com estudos anteriores, como o de Zilles (2007), que encontrou 72% de realização da variante *a gente*, na amostra do VARSUL, os resultados se apresentam conforme o esperado e mostram um aumento significativo na realização do pronome *a gente* (65% vs. 87%) nas duas amostras aqui comparadas. Além disso, se observa a inserção em um contexto de mudança generalizada nas variedades de português brasileiro, uma vez que os achados de Vitória (2017) também apresentam um elevado percentual de realização de *a gente* (84%), na função de sujeito, na fala maceioense.

Acredita-se que este estudo contribua não só com a apresentação de como se encontra a variação entre *nós* e *a gente*, na função de sujeito, na capital dos gaúchos, em um intervalo de 25 anos, mas também com um esforço inicial de inserir o fator *Estilo* na análise, discussão que será mais bem desenvolvida em etapas de pesquisa futuras e mais robustas. Espera-se que os resultados aqui apresentados sejam úteis na realização de outras investigações acerca da variação entre *nós* e *a gente*, em Porto Alegre, bem como comparáveis com outros estudos em todo o território nacional.

REFERÊNCIAS

- BAUGH, J. **A dissection of style-shifting**. In: ECKERT, P.; RICKFORD, J. R. (Eds.). *Style and sociolinguistic variation*. Cambridge: Cambridge University Press, 2001. p. 109-118.
- BATTISTI, E. **O português falado no Rio Grande do Sul: história e variação linguística**. In: BISOL, L.; BATTISTI, E. *O português falado no Rio Grande do Sul*. Porto Alegre: Edipucrs, 2014, p. 9-17
- BELL, A. **Back in style: reworking audience design**. In: ECKERT, P.; RICKFORD, J. (Eds.). *Style and sociolinguistic variation*. Cambridge: Cambridge University Press, 2001. p. 139-169.
- BORGES, P. R. S. **A gramaticalização de “a gente” no português brasileiro**. Tese (Doutorado em Letras) – Programa de Pós-Graduação em Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2004.
- CAETANO, M. M.; BOMFIM, Eneida. **Gramaticalização - de Meillet aos dia contemporâneos: parâmetros para uma pesquisa sob perspectiva pancrônica**. 2011. 107 p. Dissertação (Mestrado)-Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Letras, 2011.
- CARVALHO, M. A. S. **Uso variável de nós e a gente na fala e escrita de pessoas do Distrito Federal**. 2015. 23 f. Trabalho de Conclusão de Curso (Licenciatura em Letras Português) — Universidade de Brasília, Brasília, 2015.
- CASTILHO, A. T. **Unidirectionality or Multidirectionality? Some Issues on Grammaticalization**. In: XII Seminary of Functional Syntax. São Paulo: USP, 2002A.
- CHOMSKY, N. (1957). **Syntactic Structures**. Haia: Mouton.
- CHOMSKY, N. (1965). **Aspects of the Theory of Syntax**. Cambridge, Mass.: MIT Press.
- COUPLAND, N. **Style: language variation and identity**. Cambridge: Cambridge University Press, 2007. CRAIG, C. **Ways to Go in Rama: a Case Study on Polygrammaticalization**. In: TRAUOGOTT, E.C.; HEINE, B. (orgs.). *Approaches to Grammaticalization*. Amsterdam: John Benjamins, 1991, p. 455-492.
- ECKERT, P. **Linguistic variation as social practice**. Oxford: Blackwell, 2000.
- ECKERT, P. **Style and social meaning**. In: ECKERT, P.; RICKFORD, J. R. (Eds.). *Style and sociolinguistic variation*. Cambridge: Cambridge University Press, 2001. p. 119-126.
- FISCHER, L. A. **Dicionário de palavras & expressões estrangeiras**. Porto Alegre: L&PM, 2004.
- FREITAG, R. M. K.; MARTINS, M. A. R.; ARAÚJO, A.; BATTISTI, E.; COELHO, I. M. W. DA S.; SOUSA, M. D. A. F.; SILVA, R. G. DA; LIMA-LOPES, R. E. DE. **Challenges of Linguistic Data Management and Open Science**. *Cadernos de Linguística*, v. 2, n. 1, p. 01-19, 2021.

- GONÇALVES, S. C. L. *et al.* **Tratado geral sobre gramaticalização.** In: GONÇALVES, S. C. L.; LIMA-HERNANDES, M. C.; CASSEB-GALVÃO, V. C. (Org.). Introdução à gramaticalização: princípios teóricos e aplicação. p. 15-66. São Paulo: Parábola, 2007.
- GOFFMAN, E. **Footing.** In: RIBEIRO, Branca T.; GARCEZ, Pedro M. (Orgs.) Sociolinguística Interacional. 2.ed. São Paulo: Edições Loyola, 2002.p. 107-148.
- HEINE, B.; KUTEVA, T. **World lexicon of grammaticalization.** Cambridge: Cambridge University Press, 2002.
- HOPPER, P.; TRAUGOTT, E. **Grammaticalization.** Cambridge: Cambridge University Press, 1993.
- HEINE, B. **Grammaticalization.** In: JOSEPH, Brian; JANDA, Richard D. (Ed.). The handbook of historical linguistics. Oxford: Blackwell, 2003. p. 575-601.
- KURYŁOWICZ, J. **The evolution of grammatical categories.** Diogenes 51: 55-71. Reprint: KURYŁOWICZ, J. 1975, Esquisses linguistique II. München: W: Fink, 1965.
- LABOV, W. **Principles of linguistic change – Volume 1: Internal factors.** Oxford: Blackwell, 1994.
- LABOV, W. **Some further steps in narrative analysis.** *Journal of Narrative and Life History.* 7 (1-4): 395 – 413, 1997.
- LABOV, W. **The anatomy of style-shifting.** In: ECKERT, P.; RICKFORD, J. R. (Eds.). Style and sociolinguistic variation. Cambridge: Cambridge University Press, 2001. p. 85-108.
- LABOV, W.; WALETZKY, T. **Narrative Analysis. Oral Versions of personal experiences.** In: PAULSTON, C. B.; TUCKER, G. R. *Sociolinguistics.* The essential readings. 1. ed. London: Blackwell, 2003. p. 74-104.
- LABOV, W. **Padrões sociolinguísticos.** Trad. Marcos Bagno, Maria Marta Pereira Scherre e Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008.
- LOPES, C. R. S. **O percurso de a gente em tempo real de longa duração.** In: Para a história do português brasileiro. São Paulo: Humanitas Publicações/ FFLCH/USP, 2001. v. II, tomo I – Primeiros estudos. p. 127-148.
- MARCUSCHI, L. A. **Análise da conversação.** 5. Ed. São Paulo: Editora Ática, 2003.
- MEILLET, A. **L'évolution des formes grammaticales.** In: Linguistique historique et linguistique générale. Paris: Champion, 1912. p. 130-148.
- NEWMAYER, F.J. **Deconstructing Grammaticalization.** Language Sciences, v. 23, 2001, p. 187-299.
- SAUSSURE, F. de. **Curso de lingüística geral.** Tradução de Antônio Chelini et al. 22. ed. São Paulo: Cultrix, 2000.
- SCHILLING-ESTES, N. **Stylistic variation and the sociolinguistic interview: a reconsideration.** In: MONROY, R.; SÁNCHEZ, A. (Eds.). Actas del XXV Congreso internacional de la Asociación Española de Lingüística Aplicada (AESLA). Murcia, 2007. p.

971-986. Disponível em: <https://www.um.es/lacell/aesla/contenido/pdf/9/schilling.pdf> Acesso em: 25 jul. 2021.

SIMÕES, L. J. **Aprendizagem da gramática do português escrito: algumas reflexões a partir da língua falada.** *Calidoscópico*, v. 4, n. 1, p. 51-59, 2006.

VALLE, C. R. N., GÖRSKI, E. M. **Por um tratamento multidimensional da variação estilística na entrevista sociolinguística.** In: GÖRSKI E. M.; COELHO, I. L.; SOUZA, C. M. N. (Orgs.). *Variação estilística: reflexões teórico-metodológicas e propostas de análise.* Coleção Linguística. V.3. Florianópolis: Insular, 2014. p. 93-121.

VITÓRIO, E. G. DE S. L. A. **A variação nós e a gente na posição de sujeito na fala de crianças da cidade de Maceió/AL.** *Revista (Con)Textos Linguísticos*, v. 9, n.14, p. 126-141, 2015.

VITÓRIO, E. G. DE S. L. A. **A realização dos pronomes nós e a gente na função de sujeito e nas funções de complemento e adjunto na cidade de Maceió/AL.** *Letrônica*, v. 10, n. 1, p. 122-138, 27 dez. 2017.

VOTRE, S. **Relevância da variável escolaridade.** In: MOLLICA, Maria; BRAGA, Maria (Org.). *Introdução à Sociolinguística: o tratamento da variação.* São Paulo: Contexto, 2003. p. 51-57.

WEINREICH, U.; LABOV, W.; HERZOG, I. **Fundamentos empíricos para uma teoria da mudança linguística.** Tradução Marcos Bagno. São Paulo: Parábola Editorial, 2006.

ZILLES, A. M. S.; MAYA, L. Z.; SILVA, K. Q. **A concordância verbal com a primeira pessoa do plural em Panambi e Porto Alegre, RS.** *Organon*, 14, n. 28-29, p. 195-219, 2000.

ZILLES, A. M. S. **The development of a new pronoun: the linguistic and social embedding of a gente in Brazilian Portuguese.** *Language Variation and Change*, v. 17, n. 1, p. 19-53, 2005

ZILLES, A. M. S. **O que a fala e a escrita nos dizem sobre a avaliação social do uso de a gente?** *Letras de Hoje*, Porto Alegre, v. 42, n. 2, p. 27-44, jun. 2007.