

## VARIAÇÃO NA CONCORDÂNCIA NOMINAL DE NÚMERO NO PORTUGUÊS DE PORTO ALEGRE (RS) E OS ESTILOS DE VIDA DOS PORTO-ALEGRENSES

VARIATION IN THE NOMINAL NUMBER AGREEMENT IN PORTO ALEGRE PORTUGUESE AND THE LIFESTYLES OF PORTO-ALEGRENSES

Bruna Silva dos Santos | [Lattes](#) | [brunacortezibc@gmail.com](mailto:brunacortezibc@gmail.com)  
Universidade Federal do Rio Grande do Sul

Elisa Battisti | [Lattes](#) | [battisti.elisa@gmail.com](mailto:battisti.elisa@gmail.com)  
Universidade Federal do Rio Grande do Sul

**Resumo:** O artigo apresenta os resultados de um estudo acerca da concordância nominal de número (CN) no português em Porto Alegre (RS) no que se refere à variação entre CN explícita (*as coisas quebradas*) e CN zero (*as coisas quebrada, as coisa quebrada*). Na primeira etapa do estudo, realizou-se análise quantitativa de produção linguística do tipo atomística (Scherre, 1988) ou mórfica (Lucchesi; Dália, 2022) na linha sociolinguística variacionista (Labov, 2008), com dados de 32 entrevistas sociolinguísticas do LínguaPOA (2015-2019). Tomando-se a CN explícita como valor de aplicação da variável resposta, controlaram-se dez variáveis previsoras: Processos morfofonológicos para a formação de plural, Tonicidade do item lexical singular, Posição relativa do elemento ao núcleo, Posição Linear do elemento no sintagma nominal, Classe gramatical do item analisado, Gênero, Escolaridade, Renda, Zona. A análise dos dados com o R, uma linguagem de programação e ambiente de computação estatística (The R Core Team, 2022), em modelos de regressão logística de efeitos mistos mostrou que a CN explícita é favorecida por elementos pospostos ao núcleo do sintagma nominal e por falantes com nível superior de escolaridade. É desfavorecida por plural regular, vocábulos paroxítonos e proparoxítonos e falantes de renda baixa. Na segunda etapa do estudo, realizou-se análise de conteúdo (Bardin, 2011) em 8 das 32 entrevistas sociolinguísticas, procurando esclarecer por que categorias como Escolaridade, isoladamente, não são capazes de explicar a CN explícita. Os relatos nas entrevistas mostram que produzir mais CN explícita é atributo de sujeitos cujos estilos de vida pautam-se pelo gosto de luxo (Bourdieu, 2015): circulam pelo centro da cidade, consomem produtos culturais, em práticas que valorizam e requerem o uso da CN explícita, o que é possível com maiores níveis de renda.

**Palavras-chave:** Concordância nominal de número; Português brasileiro de Porto Alegre; Análise variacionista laboviana; Análise de Conteúdo; Dados do LínguaPOA.

**Abstract:** The article presents the results of a study about nominal number agreement (NC) in Porto Alegre Portuguese regarding the variation between explicit NC (*as coisas quebradas* ‘the broken things’) and zero NC (*as coisas quebrada, as coisa quebrada*). In the first stage of the study, a quantitative analysis of linguistic production of the atomistic (Scherre, 1988) or morphic type (Lucchesi; Dália, 2022) was carried out in the variationist sociolinguistic line (Labov, 2008), with data from 32 sociolinguistic interviews from LínguaPOA (2015-2019). Having explicit NC as the application value of the response variable, ten predictor variables were controlled: Morphophonological processes for plural formation, Stress of the singular lexical item, Relative position of the element to the head, Linear Position of the element in the noun phrase, Grammatical class of the analyzed item, Gender, Education, Income, Zone. Data analysis with R, a language and environment for statistical computing and graphics (The R Core Team, 2022), in mixed effects logistic regression models showed that explicit NC is favored by elements postponed to the head of the noun phrase and by speakers with a higher education level. It is disfavored by regular plurals, paroxytone and proparoxytone words, and low-income speakers. In the second stage of the study, content analysis was carried out (Bardin, 2011) in 8 of the 32 sociolinguistic interviews, seeking to clarify why categories such as Education, alone, are not able to explain the explicit NC. The reports in the interviews show that producing explicit NC is an attribute of subjects whose lifestyles are guided by a taste for luxury (Bourdieu, 2015): they circulate in the city center, consume cultural products, in practices that value and require the use of explicit NC, which is possible with higher income levels.

**Key-words:** Nominal number agreement; Brazilian Portuguese from Porto Alegre; Labovian variacionist analysis; Content analysis; LínguaPOA data.

## Introdução

Neste artigo, sobre a concordância nominal de número em sintagmas nominais (CN) no português de Porto Alegre,<sup>1</sup> tratamos da variação entre CN explícita (*aquelas meninas bonitas*) e CN zero (*aquelas meninas bonita*Ø, *aquelas menina*Ø *bonita*Ø). Nossos objetivos são apresentar e discutir os resultados das análises quantitativa e qualitativa realizadas,<sup>2</sup> a primeira para esclarecer os condicionadores linguísticos e sociais da CN variável em Porto Alegre (POA), a segunda, para levantar relatos de práticas sociais

<sup>1</sup> Porto Alegre é a capital do Rio Grande do Sul, estado situado no extremo sul do Brasil.

<sup>2</sup> Este artigo é um recorte de Santos (2022).

associadas ao valor social e aos eventuais significados sociais indexados pela CN explícita e pela CN zero nessa variedade de fala.

Mesmo havendo uma longa tradição de estudos variacionistas acerca da CN no português brasileiro (PB), POA carece de uma pesquisa sociolinguística para além da realizada por Pontes (1979), que investigou a variável na fala de indivíduos não alfabetizados, com ensino primário incompleto. Salientamos aqui a falta especialmente de um estudo sociolinguístico variacionista representativo da comunidade de fala, efetuado com dados mais recentes, uma vez que Mangabeira (2016) investigou a CN no português de POA, mas em uma comunidade de prática.

Assim, com este estudo, esperamos colaborar no mapeamento da CN variável no PB, a partir da análise de produção linguística na perspectiva atomística ou mórfica,<sup>3</sup> na linha variacionista quantitativa laboviana (Labov, 1994, 2001, 2008, 2010), de uma de suas variedades, o PB de POA. Considerando os resultados de investigações já realizadas acerca da CN no PB (Scherre, 1988, Oushiro, 2015, Mangabeira, 2016, entre muitos outros), nossa hipótese é a de que os condicionadores da CN variável no PB de POA sejam similares aos verificados em outras variedades, pois, como afirma Scherre (1994, p. 2), a CN apresenta “tendências sistemáticas de variação altamente previsíveis”.

Para isso, extraímos dados de 32 entrevistas sociolinguísticas do acervo LínguaPOA (2015-2019) e os analisamos quantitativamente com a linguagem de programação R (The R Core Team, 2022), na interface RStudio, em modelos estatísticos de regressão logística de efeitos mistos (função *glmer*) contendo as variáveis aleatórias Palavra e Informante. A novidade da análise quantitativa encontra-se nesse tratamento estatístico dos dados<sup>4</sup>: a execução da análise em modelos efeitos mistos, com a inclusão de variáveis aleatórias,

<sup>3</sup> Pelo recorte efetuado neste artigo, trazemos apenas a análise da CN variável na perspectiva *atomística* (Scherre, 1988) ou *mórfica* (Lucchesi; Dália, 2022). Os estudos variacionistas sobre a CN variável a partir de Scherre (1988) reconhecem dois tipos de análise, *não-atomística*, levando em conta a presença/ausência de concordância nominal de número no sintagma como um todo, e *atomística*, considerando a presença de marca de concordância em cada elemento do sintagma. Lucchesi e Dália (2022) chamam a análise não-atomística de *sintagmática* e a atomística de *mórfica*, denominação que adotamos aqui. Segundo esses autores, os estudos de Scherre e os subsequentes, de diferentes autores, acabaram por focalizar preponderantemente a perspectiva *atomística/mórfica*, o que não fornece um retrato completo da CN variável. Concordamos com os autores. A decisão de relatar aqui apenas os resultados da análise mórfica deve-se, de um lado, ao necessário recorte de informações requerido pelo artigo; de outro, à preponderância acima afirmada e, com ela, a possibilidade de comparar nossos resultados aos de vários outros estudos. Para os resultados da análise sintagmática da CN variável no PB de POA, ver Santos (2022).

<sup>4</sup> As análises quantitativas da CN variável no PB, a partir dos anos 1970, realizaram-se com os programas do pacote VARBRUL (Cedergren; Sankoff, 1974) e, depois disso, com versões desses programas para o ambiente Windows, como a Goldvarb-X (Sankoff; Tagliamonte; Smith, 2005). Análises mais recentes, como a de Oushiro (2015), Santos (2022) e a que se realiza neste artigo empregam a linguagem de programação R para a análise estatística de regressão logística, especialmente porque o R possibilita obter modelos de efeitos mistos, o que está além das potencialidades dos programas do pacote VARBRUL e Goldvarb-X.

minimiza os efeitos estatísticos de contextos linguísticos repetidos e de características específicas do informante.

Outra contribuição deste estudo está no esclarecimento de práticas sociais peculiares a categorias correlacionadas à CN variável, o que pode esclarecer as motivações sociais do processo. A CN variável é fortemente influenciada por variáveis sociais (Scherre, 1998a), sendo que a não realização de concordância pode ser socialmente estigmatizada, associada a pouca escolaridade e baixa renda (Mangabeira, 2016).<sup>5</sup> Para isso, realizamos análise de conteúdo (Bardin, 2011) das entrevistas sociolinguísticas e mobilizamos a teoria social de Bourdieu (1996, 2015) para, a partir de relatos sobre atividades cotidianas, explorar os estilos de vida dos informantes e sua possível relação com a CN variável.

Explorar estilos de vida, como fez Oliveira (2018) no estudo dos significados sociais do *ingliding* de vogais tônicas no PB de POA, é uma forma de esclarecer a relação entre variação linguística e classes sociais, investigando práticas sociais realizadas por informantes com níveis distintos de escolaridade e renda, práticas essas que estão na base de diferenças linguísticas e sociais. Nossa hipótese é a de que há uma associação das variantes CN explícita e CN zero com o contraste centro-periferia no espaço social de Porto Alegre. Veremos que, além da escolaridade e renda, práticas sociais como as de lazer, realizadas pelos informantes e provavelmente relacionadas a diferentes graus de mobilidade entre zonas da cidade (Battisti; Oushiro, 2022), têm efeito sobre a realização da CN explícita no PB de POA.

### **A concordância nominal de número no português brasileiro**

Estruturalmente, o português conta com o mecanismo da flexão, que possibilita a um vocábulo modificar-se pela relação lógico-gramatical com os demais constituintes do sintagma. A flexão em português se dá por afixação (adição de constituintes mórficos) ao final dos vocábulos (sufixação), de forma sistemática e coerente (Camara Jr., 2015), e serve ao processo de concordância, que pode implicar a repetição dos morfemas (ou alo-morfes) flexionais ao longo dos sintagmas. Além disso, a concordância também é caracterizada como um “princípio segundo o qual certos termos (dependentes, determinantes)

---

<sup>5</sup> A concepção de variante estigmatizada que mobilizamos conforma-se à Labov (2008, p. 212): “Sob extrema estigmatização, uma forma se torna assunto de comentário social explícito e pode acabar por desaparecer. Trata-se então de um *estereótipo* que pode ficar cada vez mais divorciado das formas que são realmente usadas na fala”. Porém, diferentemente do que prevê Labov, a CN zero não desapareceu, dá mostras de estabilidade no sistema do PB, ensejando discussões como a do caso, de 2011, do livro didático *Para uma vida melhor*, que apresentou uma seção sobre variação linguística e abordou a questão da concordância nominal e verbal. Na época, o contexto “*os livros*” foi alvo de muitos comentários puristas acerca do que seria uma língua correta, certa, legítima. Para mais informações sobre o assunto, recomendamos a leitura de Cavalcanti (2013) e Lucchesi (2011).

se adaptam, na forma, às categorias gramaticais de outros (principais, determinados)” (Luft, 2002, p. 42), o que exige a marcação redundante da flexão nos elementos flexionáveis do sintagma.

Todavia, estudos sociolinguísticos acerca da CN variável na fala vernacular do PB (Scherre, 1988, Oushiro, 2015, Mangabeira, 2016, entre outros) revelam que os elementos flexionáveis nem sempre concordam com o núcleo. O primeiro elemento do sintagma recebe categoricamente a marca de plural, sustentando o sentido pluralizado do constituinte, razão pela qual o sintagma nominal pode não ser redundantemente flexionado para número. Logo, na expressão oralizada do PB, a CN se caracteriza como uma regra variável em função de os falantes ora realizarem (*os pilas*), ora não realizarem a concordância de número (*os pilaØ*) redundante ao longo do sintagma.

Assim, na fala, pode-se observar marcação de número nos elementos flexionáveis do sintagma – CN *explícita*, nos termos de Scherre (1988), Oushiro (2015) – ou ausência de marcação em pelo menos um dos elementos do sintagma – CN *zero* em um ou mais elementos, conforme essas mesmas autores. No excerto (1), observam-se instanciações de marcação de plural tanto explícita (nos trechos em negrito) quanto zero (registrada com Ø nos trechos em itálico) em um mesmo enunciado, produzido por um mesmo informante (variação intraindividual).

- (1) É eu ã eu trabalhei tipo faz uns cinco seis meses mais ou menos que eu saí da (nome da empresa) né trabalhei **quinze anos** na (nome da empresa) de distribuição de jornal né e aí tipo não deu pra conciliar tipo *esses quinze anoØ* que eu trabalhei na (nome da empresa) eh não rendeu aqui pra ONG assim né porque tipo trabalhava de noite tinha que fazer reuniões **essas coisas** durante o dia né e agora com *os filhoØ tudo criadoØ* aí eu pensei agora vou me dedicar mesmo à ONG aqui me dedicar ao trabalho voluntário né e tamo aí na luta né fazendo *esses trabalhoØ comunitárioØ* e **nos fins de semana** a gente faz festa de aniversário casamento é assim que a gente vai tirando pra sobreviver. (LÍNGUAPOA, 2015-2019, Inf97, grifos nossos).<sup>6</sup>

Cabe salientar que a CN variável é amplamente investigada no Brasil e, mesmo em comunidades de fala distintas, apresenta muitas similaridades quanto aos condicionado-

<sup>6</sup> O fragmento é da entrevista de um informante de ensino básico e baixa renda, morador do bairro Mario Quintana (ver figura 1). Localizado na zona leste de Porto Alegre, esse bairro é considerado um dos mais pobres da cidade (Cf. ObservaPOA, disponível em <http://www.observapoa.com.br/default.php>, acesso em 10/11/2022).



res e ao encaixamento. A questão suscitada pelo conjunto de estudos é se as similaridades nos resultados das análises devem-se ao fato de a CN se apresentar como uma variável estável nos diferentes recortes sincrônicos, com poucas divergências nas correlações atestadas, como aponta Scherre (1994, p. 2, grifos nossos):

O fenômeno da variação na concordância de número no português falado do Brasil, longe de ser restrito a uma região ou classe social específica, é *característico de toda a comunidade de fala brasileira, apresentando diferenças mais de grau do que de princípio* [...] o fenômeno de variação de número no português do Brasil pode ser caracterizado como um caso de variação linguística inerente, tendo em vista que *ocorre em contextos linguísticos e sociais semelhantes e apresenta tendências sistemáticas de variação altamente previsíveis*. (Scherre, 1994, p. 2, grifos nossos)

Ou se, para manter a tradição de análise de Scherre (1988), os estudos posteriores se preocuparam mais com atestar os resultados da pesquisadora do que com propor avanços metodológicos, como sugerem Lucchesi e Dália (2022):

A variação na concordância nominal de número é um dos aspectos da morfossintaxe do português mais estudados pela Sociolinguística Variacionista no Brasil, e a tese de Marta Scherre (1988), intitulada *Reanálise da Concordância Nominal em Português*, constituiu um marco no desenvolvimento das análises variacionistas do fenômeno. Scherre destaca como principais condicionamentos estruturais da variação na concordância nominal de número a posição do constituinte e sua natureza (nuclear ou não), a diferença entre a forma do singular e do plural do constituinte, com base no princípio da saliência fônica, e a presença ou ausência de marca de plural no constituinte precedente, no que se denominou paralelismo formal. Desde então, praticamente todas as análises do fenômeno seguem esse enquadramento [...] Como a maioria desses estudos se contentou em comprovar os achados de Scherre (1988), poucos avanços foram alcançados na compreensão de como o mecanismo da concordância nominal de número é condicionado na estrutura da língua. Assim, algumas lacunas e incompreensões se mantêm, sobretudo na forma de abordar o fenômeno. (Lucchesi; Dália, 2022, p. 7370).

Mesmo reconhecendo as críticas de Lucchesi e Dália (2022), entendemos que Marta Scherre iniciou uma tradição nos estudos sociolinguísticos acerca da CN no Brasil, produzindo conhecimentos que ainda vêm sendo ampliados. Inserindo-nos nessa tradição, e pelas limitações da divulgação do trabalho em formato de artigo, optamos por divulgar apenas os resultados na análise mórfica, e parte dos resultados da análise qualitativa.

## Metodologia

A análise quantitativa aqui realizada, de produção linguística da CN variável no PB de POA, orienta-se pela sociolinguística variacionista (Labov, 2008), que se pauta pela ideia de a heterogeneidade linguística manifesta na fala ser ordenada, correlacionada a variáveis sociais e linguísticas.

Como afirmamos na Introdução, os dados extraídos de 32 entrevistas sociolinguísticas do acervo LínguaPOA (2015-2019) são analisados quantitativamente na perspectiva mórfica. Usa-se, para tanto, a linguagem de programação R (The R Core Team, 2022) na interface RStudio. Os 32 informantes de cujas entrevistas extraíram-se os dados têm entre 40 e 59 anos e são equilibradamente estratificados em Zona (centro, leste, sul, norte), Gênero (masculino, feminino), Escolaridade (Básico, Superior), Renda média mensal dos domicílios no bairro de residência do informante (baixa, alta),<sup>7</sup> conforme mostra o quadro 1.

**Quadro 1.** Estratificação dos informantes do LínguaPOA (2015-2019) considerados na análise. Os números nas células do quadro identificam cada informante no LínguaPO

|                 | Centro |     |      |     | Leste |     |      |     | Sul |     |      |     | Norte |     |      |     |
|-----------------|--------|-----|------|-----|-------|-----|------|-----|-----|-----|------|-----|-------|-----|------|-----|
|                 | Fem    |     | Masc |     | Fem   |     | Masc |     | Fem |     | Masc |     | Fem   |     | Masc |     |
| Escol.          | Bás    | Sup | Bás  | Sup | Bás   | Sup | Bás  | Sup | Bás | Sup | Bás  | Sup | Bás   | Sup | Bás  | Sup |
| Renda*<br>baixa | 29     | 30  | 26   | 27  | 100   | 102 | 97   | 99  | 137 | 138 | 134  | 135 | 65    | 66  | 62   | 63  |
| Renda*<br>alta  | 10     | 12  | 8    | 9   | 83    | 84  | 80   | 81  | 119 | 120 | 116  | 117 | 47    | 48  | 44   | 45  |

\* Renda média mensal dos domicílios no bairro de residência do informante.

Informantes em azul têm nível Fundamental de escolaridade e foram agrupados ao nível Médio na categoria Básico.

Fonte: Elaborado pelas autoras.

Na análise estatística, a CN foi tratada como variável binomial, compreendendo CN explícita e CN zero. Sendo uma análise do tipo mórfica, os elementos dentro dos sintagmas são computados individualmente. Por exemplo, em um sintagma como *esses trabalhoØ comunitárioØ*, há duas ocorrências de CN zero e uma ocorrência de CN explícita. O valor de aplicação da variável resposta é a CN explícita. As variáveis previsoras linguísticas, baseadas na literatura revisada – especialmente Scherre (1988), Oushiro (2015), Mangabeira (2016) – são: *Processos morfofonológicos* para a formação de plural e

<sup>7</sup> As quatro zonas, os dois gêneros, os níveis de escolaridade e renda considerados seguem os critérios de estratificação do LínguaPOA. Informações sobre esses critérios estão disponíveis em <https://www.ufrgs.br/linguapoa/> Acesso em 10 nov. 2022.

*Tonicidade* do item lexical singular, como variáveis separadas, que contemplam aspectos abrangidos pela variável *Saliência Fônica* (cf. SCHERRE, 1988); *Posição linear* do elemento no sintagma nominal; *Posição Relativa* do elemento ao núcleo; *Classe gramatical* do item analisado. As variáveis predictoras sociais são: *Gênero*, apontada por Scherre (1998) como uma das variáveis geralmente correlacionadas à CN variável; *Escolaridade*, uma vez que, conforme Scherre (1988, 1998a), a proporção de CN explícita é proporcional ao nível de escolaridade (quanto mais alto o nível de escolaridade, mais uso da marca explícita de plural); *Renda*, variável relacionada ao acesso a certas práticas sociais; e *Zona* de residência dos falantes, já que as práticas sociais nelas e entre elas realizadas influenciam os estilos de vida e os padrões de mobilidade no espaço urbano dos informantes (Oliveira, 2018, AUTOR; COAUTOR, 2022), o que pode se correlacionar à CN variável. No quadro 2, estão dispostas as variáveis predictoras investigadas e seus respectivos níveis.

**Quadro 2.** Variáveis predictoras

| Variável                                             | Fatores                                                                                |
|------------------------------------------------------|----------------------------------------------------------------------------------------|
| Processos morfofonológicos para a formação de plural | Plural regular<br>Plural irregular                                                     |
| Tonicidade do item lexical singular                  | Monossílabos tônicos e oxítonos<br>Monossílabos átonos e paroxítonos<br>Proparoxítonos |
| Posição relativa do elemento ao núcleo               | Pré-nuclear<br>Núcleo<br>Pós-nuclear                                                   |
| Posição Linear do elemento no Sintagma Nominal       | Primeira posição<br>Segunda posição<br>Terceira posição<br>Outras posições             |
| Classe gramatical do item analisado                  | Elemento nominal<br>Elemento não nominal                                               |
| Gênero                                               | Feminino<br>Masculino                                                                  |
| Escolaridade                                         | Básico (Fundamental+Médio)<br>Superior                                                 |
| Renda                                                | Baixa<br>Alta                                                                          |



|      |                                 |
|------|---------------------------------|
| Zona | Centro<br>Leste<br>Sul<br>Norte |
|------|---------------------------------|

Fonte: Elaborado pelas autoras.

Além dessas variáveis, adicionam-se aos modelos de regressão logística de efeitos mistos as variáveis aleatórias Informante e Palavra.

Em sua tese de doutoramento, Scherre (1988) contempla a variável Saliência Fônica<sup>8</sup> desdobrando-a em duas variáveis: *Processos morfofonológicos* de formação de plural<sup>9</sup> e *Tonicidade do item singular*<sup>10</sup>, o que também fazemos aqui. Assumindo, como Scherre (1989), que os efeitos de Processos e Tonicidade se sobrepõem – há “uma forte sobreposição entre as duas variáveis, decorrente da realidade linguística que envolve os dados analisados” (Scherre, 1989, p. 306) – são feitas, no presente estudo, análises estatísticas conforme o sugerido por Scherre (1988): além do modelo com ambas as variáveis, testamos outros dois. Retiramos de um deles a variável Processos, de outro, a variável Tonicidade. Nosso objetivo é verificar os efeitos, sobre as estimativas, da presença-ausência de Processos e Tonicidade nos modelos, uma vez que “inevitavelmente haverá diferenças probabilísticas” entre eles (Scherre, 1989, p. 307). Fechando o conjunto de variáveis linguísticas, o estudo contou com as variáveis *Posição linear*, *Posição relativa* e *Classe gramatical*. Por sua não ortogonalidade, também planejavamos analisá-las separadamente. No entanto, como veremos no relato dos resultados, essa medida não foi necessária porque as variáveis *Posição linear* e *Classe gramatical* acabaram não sendo

<sup>8</sup> A variável Saliência Fônica é frequentemente considerada em estudos acerca da CN variável. Ela foi inicialmente elaborada por Naro e Lemle (1977 *apud* Chaves, 2014). Resumidamente, Saliência Fônica diz respeito à diferenciação do material fônico dos elementos linguísticos no contraste singular/plural: a marca de plural em *casa/casas*, por exemplo, é menos saliente do que a marca em *lençol/lençóis*.

<sup>9</sup> Uma das diferenças de nossa análise em relação à de Scherre (1988) é o número de níveis na variável Processos. A autora controla seis níveis (plural metafônico, itens com final em /l/, em /ão/, em /r/, em /s/ e plurais regulares) distribuídos em um *continuum* de saliência. De um lado, têm-se os elementos menos salientes (plurais regulares) e, de outro, observam-se os mais salientes (metafônicos). Também fizemos isso neste trabalho, porém, devido ao número desequilibrado de dados por nível, optamos por amalgamar os seis níveis em dois: plurais regulares e plurais irregulares. Além disso, observamos que a maioria dos elementos com plurais irregulares são oxítonos, o que desequilibra o número de plurais irregulares paroxítonos e proparoxítonos na amostra, impossibilitando a análise. Com isso, decidimos manter a diferenciação entre plurais regulares e irregulares, mas apenas os plurais regulares apresentaram diferenciação quanto à sua Tonicidade (cf. Quadro 2).

<sup>10</sup> Para ser mais exato, Scherre (1988) investigou uma terceira dimensão da Saliência Fônica: Número de Sílabas. A variável, porém, não se mostrou relevante para o estudo da CN.

incluídas nos modelos devido a problemas nos níveis de cada uma delas – CN explícita quase categórica em um nível de *Posição linear*, ausência de diferença significativa entre os níveis de *Classe gramatical*.

As hipóteses testadas no controle das variáveis previsoras linguísticas estão de acordo com a literatura. Para *Processos*, seguimos Scherre (1988), Oushiro (2015) e Mangabeira (2016) e esperamos observar um favorecimento de CN explícita em plurais irregulares (*os lençóis*, *meus irmãos*), desfavorecimento em plurais regulares (*os gêmeos*, *as casas*). Para *Tonicidade*, a expectativa é de que itens lexicais paroxítonos e proparoxítonos (*os cara*, *os médico*) desfavoreçam a CN explícita. Para *Posição Relativa*, apoiamo-nos em Oushiro (2015) e contamos que a posição nuclear (*essas atividades*) desfavoreça a marca explícita de plural, em contraste com os elementos pré-nucleares (*nosso sobrinhos*), que, esperamos, a favoreçam.

Para as variáveis previsoras sociais *Gênero*, *Escolaridade* e *Renda*, baseando-nos em Scherre (1988) e Oushiro (2015), esperamos observar um favorecimento da CN explícita por informantes do gênero feminino, com educação superior e de renda alta. Já para a variável *Zona*, esperamos que moradores das zonas Leste, Sul, Norte apresentem maior proporção de CN zero, em contraste com os residentes na zona Centro, com mais CN explícita. Nossa hipótese é a de que o Centro de POA mostre-se mais conservador porque sua formação é mais antiga do que as outras, além de ser onde situam-se instituições culturais e administrativas da cidade e do estado.

Já a análise qualitativa fundamenta-se na teoria social de Bourdieu (1996, 2015). A partir dela, procuramos esclarecer as práticas sociais de sujeitos de diferentes níveis de escolaridade, renda e local de residência enquanto estilos de vida pautados por gostos distintos, os quais promoveriam a CN explícita ou sua realização zero. Nossa hipótese é a de que renda e escolaridade, na base dos capitais econômico e cultural dos sujeitos, definem diferentes estilos de vida. Pessoas com escolaridade mais alta, mais bem situadas no mercado de trabalho, têm maiores salários, o que oportuniza o acesso a locais de relativo privilégio social, onde a marca explícita de CN é eventualmente prestigiada. Além disso, esperamos ser capazes de examinar o efeito da relação centro-periferia na CN variável: de um lado, os informantes residentes no centro da cidade apresentariam maior proporção de CN explícita; por outro, os informantes de áreas mais periféricas apresentariam mais CN zero, a não ser quando tivessem de se locomover para o centro para realizar práticas como trabalho e lazer. Em outras palavras, os informantes da periferia com maior mobili-

dade geográfica e deslocamento para o centro, onde, acreditamos, a CN explícita é relativamente prestigiada, poderiam apresentar maior proporção de concordância.<sup>11</sup>

Para testar tais hipóteses, voltamo-nos às entrevistas sociolinguísticas e às fichas sociais dos informantes. Nas entrevistas, buscamos relatos das práticas sociais dos informantes, nas fichas, informações adicionais sobre suas categorias sociais.

No exame das entrevistas, efetuamos análise de conteúdo (Bardin, 2011), técnica de estudo que objetiva a descrição do conteúdo de uma comunicação – a entrevista sociolinguística, em nosso caso. A unidade de análise é o registro temático. Consideramos, nessa análise, apenas 8 das 32 entrevistas de que, antes, extraímos dados para a análise quantitativa de produção linguística. Escolhemos 8 informantes com proporções distintas de CN zero, distribuídos ao longo de um *continuum* que vai de 0,7% a 44% de CN zero (Tabela 1).

**Tabela 1.** Informações sociais e proporção de CN zero de informantes cujas entrevistas sociolinguísticas são submetidas a análise de conteúdo

| Informante | Escolaridade | Renda* | Zona/Bairro                    | Proporção CN-parcial |
|------------|--------------|--------|--------------------------------|----------------------|
| Inf09      | Superior     | Alta   | Zona central (Menino Deus)     | 0,7%                 |
| Inf81      | Superior     | Alta   | Zona leste (Rio Branco)        | 2%                   |
| Inf117     | Superior     | Alta   | Zona sul (Tristeza)            | 7%                   |
| Inf45      | Superior     | Alta   | Zona norte (Jardim Itu Sabará) | 13%                  |
| Inf26      | Básica       | Baixa  | Zona central (Cidade Baixa)    | 14%                  |
| Inf134     | Básica       | Baixa  | Zona sul (Restinga)            | 20%                  |
| Inf97      | Básica       | Baixa  | Zona leste (Mario Quintana)    | 37%                  |
| Inf62      | Básica       | Baixa  | Zona norte (Cristo Redentor)   | 44%                  |

\* Renda média mensal dos domicílios no bairro de residência do informante

Fonte: Elaborado pelas autoras

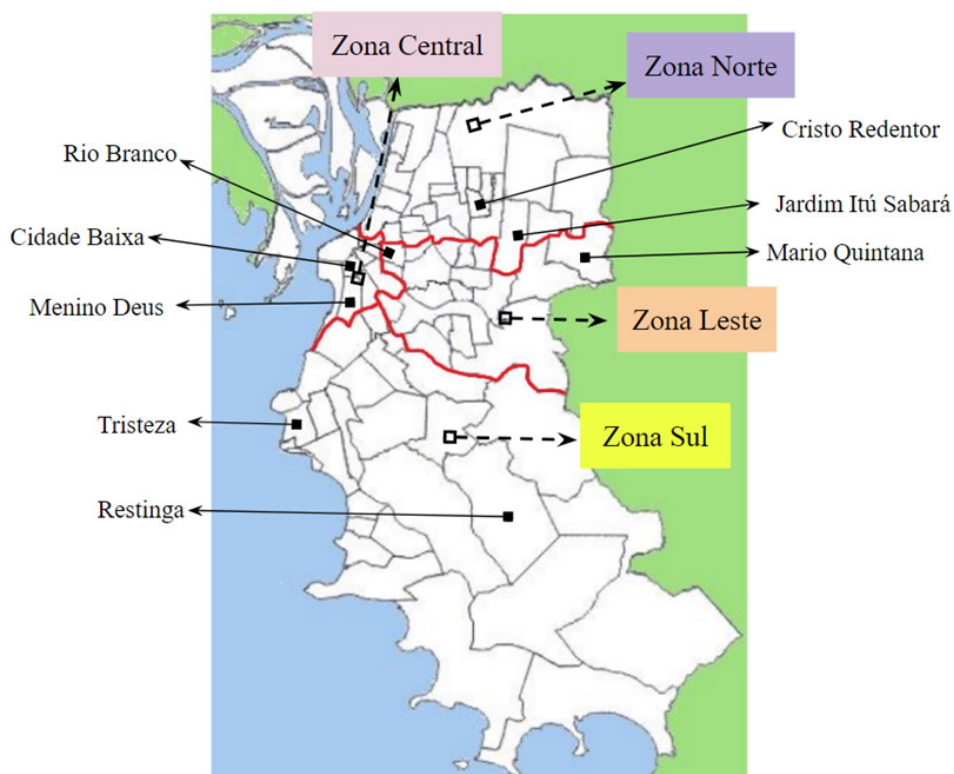
Os informantes considerados na análise qualitativa (tabela 1) distribuem-se ao longo de um *continuum* de CN zero. Essa distribuição objetivou verificar se diferenças e semelhanças na aplicação de CN zero e, por oposição, de CN explícita seriam análogas a semelhanças e diferenças das práticas sociais dos informantes. Além das proporções de CN zero no *continuum*, levamos em conta, na seleção dos 8 informantes, o fato de serem do gênero masculino, respeitando-se, também, sua distribuição por zona, em articulação com escolaridade e renda.

<sup>11</sup> A variável Mobilidade foi operacionalizada na análise quantitativa da haplologia variável por Autor e Coautor (2022), que comprovaram a correlação entre a maior mobilidade no espaço urbano e o processo variável por eles investigado. Essa operacionalização poderá ser replicada, futuramente, em análises da CN variável no PB de POA.

A escolha de entrevistas de informantes do gênero masculino para a análise de conteúdo deveu-se aos resultados da análise quantitativa de produção: a CN explícita é menos frequente na fala de informantes do gênero masculino do que do gênero feminino. Optamos pela junção de escolaridade e renda para testar a hipótese de que essas variáveis, juntas, associam-se à participação dos sujeitos em diferentes práticas sociais, viabilizadas por seus capitais e determinadas por seus *habitus*, conforme Bourdieu (2015). Contar com dois informantes de cada zona, por seu turno, voltou-se à testagem da hipótese acerca da relação entre centro-periferia com a CN.

Na Figura 1, está o mapa de Porto Alegre, com a indicação do bairro de residência de cada um dos 8 informantes considerados na análise qualitativa (tabela 1).

**Figura 1.** Localização geográfica dos 8 informantes (Tabela 1) por bairro e zona



Fonte: Elaborado pelas autoras

### **“Os mesmos desfavorecido”: os resultados da análise quantitativa**

Das 32 entrevistas sociolinguísticas do acervo LínguaPOA (2015-2019), extraímos um total de 8719 dados.<sup>12</sup> Dentre eles, 92,7% apresentam CN explícita, 7,3% apresentam

<sup>12</sup> Na análise mórfica aqui realizada, os dados são cada um dos elementos dentro do sintagma, não o sintagma como um todo.

CN zero. Ou seja, considerando-se a amostra total, a CN explícita é alta no PB de POA.

Em uma primeira etapa da análise, examinamos a distribuição dos dados nos níveis compreendidos em cada variável previsor e realizamos testes de qui-quadrado (de Pearson), para verificar se havia diferença significativa ( $p \leq 0,05$ ) nas proporções de CN explícita entre os níveis das variáveis previsoras. Incluiríamos nos modelos de regressão logística de efeitos mistos apenas as variáveis com valor de p significativo no teste de qui-quadrado. As variáveis *Processos*, *Tonicidade*, *Posição Linear*, *Posição Relativa*, *Classe Gramatical*, *Gênero*, *Escolaridade*, *Renda*, *Zona* apresentaram valor de p significativo. No entanto, constatamos que as variáveis linguísticas *Posição linear* e *Posição relativa* apresentavam um de seus níveis com quase 100% de CN explícita. Em *Posição Linear*, o nível *primeira posição* contava com um total de 3559 dados, dos quais apenas 16 apresentavam CN zero.<sup>13</sup> Por isso, optamos por retirar o nível *primeira posição* da análise. Diminuiu-se, conseqüentemente, o número total de dados (N), como se vê na tabela 2. Submeteu-se novamente a variável *Posição linear* (ajustada) ao teste de qui-quadrado, mas o valor de p (de 0,39) não se mostrou significativo, muito provavelmente porque os níveis remanescentes apresentaram praticamente a mesma proporção de CN explícita. Assim, a variável *Posição Linear* não foi incluída nos modelos.

**Tabela 2** - Distribuição de dados e aplicação da CN explícita por nível da variável *Posição linear do elemento no sintagma nominal*

|                  | N=8719 dados, <u>com</u> os dados do nível 'Primeira posição' | N=5160 dados, <u>sem</u> os dados do nível 'Primeira posição' |
|------------------|---------------------------------------------------------------|---------------------------------------------------------------|
| Níveis           | Aplicação/Tokens (%)                                          | Aplicação/Tokens (%)                                          |
| Primeira posição | 3543/3559 (99%)                                               |                                                               |
| Segunda posição  | 3728/4252 (87%)                                               | 3735/4252 (87%)                                               |
| Terceira posição | 708/791 (89%)                                                 | 708/791 (89%)                                                 |
| Outras posições  | 102/117 (87%)                                                 | 102/117 (87%)                                                 |

Fonte: Elaborada pelas autoras.

Outra variável com aplicação quase categórica de CN explícita em um dos níveis foi *Posição Relativa*: o nível *anteposta* apresentou 98% de aplicação. Porém, diferentemente de *Posição Linear*, a variável *Posição Relativa* seguiu exibindo valor de p significativo

<sup>13</sup> Esses 16 contextos são: [aqui nesta praça tem] *muito* drogados (Inf29); *o* nomes, *o* meus pais (Inf30); *a* minhas sobrinhas (Inf80); *todo* *o* meus vizinhos, *o* meus netos, *curso* técnicos (Inf119); *um* fatores positivos (Inf117); *o* efetivos (Inf62); *o* meu outros filhos, *o* meus primos, *o* meu primos, *o* meu negócios (Inf44); *o* nossos problema (Inf45); *o* meus pais (Inf47); *coisa* absurdas (Inf66).



( $p < 0,001$ ), mesmo com a retirada do nível *anteposta* da análise (tabela 3). Desse modo, optamos por efetuar as análises de regressão logística de efeitos mistos em modelos com a variável *Posição Relativa* ajustada, sem a presença do nível *anteposta* e dos dados a ele referentes. Isso causou uma redução no N, que foi de 8719 para 4942 dados.

**Tabela 3** - Distribuição de dados e aplicação da CN explícita por nível da variável *Posição relativa do elemento ao núcleo*

|           | N=8719 dados, <u>com</u> os dados do nível 'Anteposta' | N=4942 dados, <u>sem</u> os dados do nível 'Anteposta' |
|-----------|--------------------------------------------------------|--------------------------------------------------------|
| Níveis    | Aplicação/Tokens (%)                                   | Aplicação/Tokens (%)                                   |
| Anteposta | 3744/3777 (99%)                                        |                                                        |
| Núcleo    | 3887/4440 (87%)                                        | 3887/4440 (87%)                                        |
| Posposta  | 457/502 (91%)                                          | 457/502 (91%)                                          |

Fonte: Elaborada pelas autoras.

Esse ajuste em *Posição relativa*, com a retirada do nível *anteposta* e a diminuição no N de 8719 para 4942 dados, teve implicações na variável *Classe gramatical*. Ao retirar os dados referentes ao nível *anteposta*, houve redistribuição dos dados nos dois níveis de *Classe gramatical* e redução considerável no número de dados do nível *não nominal* (tabela 4). Além disso, e mais importante, os níveis apresentaram a mesma proporção de CN explícita, razão pela qual a variável não alcançou valor de  $p$  significativo, levando-nos a não incluir a variável *Classe gramatical* nos modelos de efeitos mistos.

**Tabela 4** - Distribuição de dados e aplicação da CN explícita por nível da variável *Classe gramatical do item analisado*

|                      | N=8719 dados         | N=4942 dados         |
|----------------------|----------------------|----------------------|
| Níveis               | Aplicação/Tokens (%) | Aplicação/Tokens (%) |
| Elemento nominal     | 3621/3659 (98%)      | 4323/4918 (87%)      |
| Elemento não nominal | 4460/5060 (88%)      | 21/24 (87%)          |

Fonte: Elaborada pelas autoras.

Assim, após a testagem de qui-quadrado, do exame da proporção de CN explícita por nível das variáveis e dos ajustes nos níveis, adicionamos aos modelos de análise de regressão logística de efeitos mistos apenas as variáveis *Processos*, *Tonicidade*, *Posição Relativa*, *Gênero*, *Escolaridade*, *Renda*, *Zona*.

Como esclarecemos antes (ver Metodologia), as variáveis *Processos* e *Tonicidade* foram controladas juntas em um dos modelos e separadas em outros dois, para verificar se as estimativas e correlações medidas nos modelos são afetadas. Na tabela 5 está o modelo com *Processos* e *Tonicidade*, na tabela 6 consta o modelo com *Tonicidade* e na tabela 7 vai o modelo com *Processos*.

**Tabela 5** - Análise de regressão logística de efeitos mistos da CN explícita

| N = 4942                 |                 |            |             |         |           |
|--------------------------|-----------------|------------|-------------|---------|-----------|
| Intercept = 30,2281      |                 |            |             |         |           |
| Variável                 | Aplic./Tokens   | Estimativa | Erro padrão | Valor z | p         |
| <b>Tonicidade</b>        |                 |            |             |         |           |
| Oxítone (valor de ref.)  | 1191/1271 (93%) |            |             |         |           |
| Paroxítone               | 3007/3503 (85%) | -3,074     | 0,585       | -5,250  | <0,001*** |
| Proparoxítone            | 146/168 (86%)   | -4,373     | 0,858       | -5,135  | <0,001*** |
| <b>Posição Relativa</b>  |                 |            |             |         |           |
| Núcleo                   | 3887/4440 (87%) |            |             |         |           |
| Posposta                 | 457/502 (91%)   | -2,970     | 0,440       | -7,404  | <0,001*** |
| <b>Gênero</b>            |                 |            |             |         |           |
| Feminino (valor de ref.) | 2260/2476 (91%) |            |             |         |           |
| Masculino                | 2084/2466 (84%) | -1,027     | 0,340       | -3,017  | 0,002**   |
| <b>Escolaridade</b>      |                 |            |             |         |           |
| Básica (valor de ref.)   | 1834/2200 (83%) |            |             |         |           |
| Superior                 | 2510/2742 (91%) | 1,073      | 0,343       | 3,122   | 0,001**   |
| <b>Renda</b>             |                 |            |             |         |           |
| Alta (valor de ref.)     | 2206/2405 (91%) |            |             |         |           |
| Baixa                    | 2138/2537 (84%) | -1,629     | 0,360       | -4,525  | <0,001*** |
| <b>Zona</b>              |                 |            |             |         |           |
| Centro (valor de ref.)   | 1304/1439 (90%) |            |             |         |           |
| Leste                    | 1075/1273 (84%) | -1,982     | 0,528       | -3,752  | <0,001*** |
| Norte                    | 857/985 (87%)   | -2,416     | 0,552       | -4,376  | <0,001*** |
| Sul                      | 1108/1245 (88%) | -1,723     | 0,539       | -3,194  | 0,001**   |

Modelo 1. (Aplicacao ~ Processos + Tonicidade + Posicao Relativa + Genero + Escolaridade + Renda + Zona (1|INFORMANTE) + (1|PALAVRA), data = atomistica, family = binomial)

Fonte: Elaborada pelas autoras.

Os resultados na Tabela 5 mostram que, no modelo com *Processo* e *Tonicidade* juntas, apenas as variáveis *Gênero* e *Zona* não se correlacionam à CN explícita. As estimativas nesse modelo mostram que plural regular, vocábulos paroxítonos e proparoxítonos e falantes de renda baixa desfavorecem a CN explícita. Já a posição relativa posposta ao núcleo e a escolaridade Superior favorecem a CN explícita. Esses resultados vão ao encontro de nossas hipóteses.

Retirando-se a variável *Processos* do modelo (tabela 6), todas as variáveis previsoras controladas correlacionam-se à CN explícita.

**Tabela 6. Análise de regressão logística de efeitos mistos da CN explícita sem *Processos***

| N = 4942<br>Intercept = 13,5062 |                 |            |             |         |           |
|---------------------------------|-----------------|------------|-------------|---------|-----------|
| Variável                        | Aplic./Tokens   | Estimativa | Erro padrão | Valor z | p         |
| <b>Tonicidade</b>               |                 |            |             |         |           |
| Oxítone (valor de ref.)         | 1191/1271 (93%) |            |             |         |           |
| Paroxítone                      | 3007/3503 (85%) | -3,074     | 0,585       | -5,250  | <0,001*** |
| Proparoxítone                   | 146/168 (86%)   | -4,373     | 0,858       | -5,135  | <0,001*** |
| <b>Posição Relativa</b>         |                 |            |             |         |           |
| Núcleo                          | 3887/4440 (87%) |            |             |         |           |
| Posposta                        | 457/502 (91%)   | -2,970     | 0,440       | -7,404  | <0,001*** |
| <b>Gênero</b>                   |                 |            |             |         |           |
| Feminino (valor de ref.)        | 2260/2476 (91%) |            |             |         |           |
| Masculino                       | 2084/2466 (84%) | -1,027     | 0,340       | -3,017  | 0,002**   |
| <b>Escolaridade</b>             |                 |            |             |         |           |
| Básica (valor de ref.)          | 1834/2200 (83%) |            |             |         |           |
| Superior                        | 2510/2742 (91%) | 1,073      | 0,343       | 3,122   | 0,001**   |
| <b>Renda</b>                    |                 |            |             |         |           |
| Alta (valor de ref.)            | 2206/2405 (91%) |            |             |         |           |
| Baixa                           | 2138/2537 (84%) | -1,629     | 0,360       | -4,525  | <0,001*** |
| <b>Zona</b>                     |                 |            |             |         |           |
| Centro (valor de ref.)          | 1304/1439 (90%) |            |             |         |           |
| Leste                           | 1075/1273 (84%) | -1,982     | 0,528       | -3,752  | <0,001*** |
| Norte                           | 857/985 (87%)   | -2,416     | 0,552       | -4,376  | <0,001*** |
| Sul                             | 1108/1245 (88%) | -1,723     | 0,539       | -3,194  | 0,001**   |

Modelo 2. (Aplicacao ~ Tonicidade + Posicao Relativa + Genero + Escolaridade + Renda + Zona (1|INFORMANTE) + (1|PALAVRA), data = atomistica, family = binomial)

Fonte: Elaborada pelas autoras.

Além das correlações já atestadas no modelo da tabela 5, o modelo na tabela 6 exibe correlação com as variáveis *Gênero* e *Zona*: a estimativas mostram que gênero masculino, bem como as zonas leste, norte, sul desfavorecem a CN explícita.

Já no modelo sem *Tonicidade*, mas com *Processos* (tabela 7), apenas as correlações com as variáveis linguísticas se mantêm.

Tabela 7. Análise de regressão logística de efeitos mistos da CN explícita sem Tonicidade

| N = 4942                      |                 |            |             |         |           |
|-------------------------------|-----------------|------------|-------------|---------|-----------|
| Intercept = 33,4716           |                 |            |             |         |           |
| Variável                      | Aplic./Tokens   | Estimativa | Erro padrão | Valor z | p         |
| <b>Processos</b>              |                 |            |             |         |           |
| Plural irreg. (valor de ref.) | 512/538 (95%)   |            |             |         |           |
| Plural regular                | 3832/4404 (87%) | -19,918    | 4,476       | -4,450  | <0,001*** |
| <b>Posição Relativa</b>       |                 |            |             |         |           |
| Núcleo                        | 3887/4440 (87%) |            |             |         |           |
| Posposta                      | 457/502 (91%)   | -2,783     | 0,803       | -3.464  | <0,001*** |
| <b>Gênero</b>                 |                 |            |             |         |           |
| Feminino (valor de ref.)      | 2260/2476 (91%) |            |             |         |           |
| Masculino                     | 2084/2466 (84%) | -0,977     | 0,800       | -1,222  | 0,221     |
| <b>Escolaridade</b>           |                 |            |             |         |           |
| Básica (valor de ref.)        | 1834/2200 (83%) |            |             |         |           |
| Superior                      | 2510/2742 (91%) | 1,098      | 0,775       | 1,416   | 0,156     |
| <b>Renda</b>                  |                 |            |             |         |           |
| Alta (valor de ref.)          | 2206/2405 (91%) |            |             |         |           |
| Baixa                         | 2138/2537 (84%) | -1,024     | 0,774       | -1,323  | 0,185     |
| <b>Zona</b>                   |                 |            |             |         |           |
| Centro (valor de ref.)        | 1304/1439 (90%) |            |             |         |           |
| Leste                         | 1075/1273 (84%) | -1,784     | 1,066       | -1,674  | 0,094 .   |
| Norte                         | 857/985 (87%)   | -2,041     | 1,157       | -1,763  | 0,077 .   |
| Sul                           | 1108/1245 (88%) | -1,500     | 1,124       | -1,334  | 0,182     |

Modelo 3. (Aplicacao ~ Processos + Posicao Relativa + Genero + Escolaridade + Renda + Zona (1|INFORMANTE) + (1|PALAVRA), data = atomistica, family = binomial)

Fonte: Elaborada pelas autoras.

No modelo na tabela 7, as variáveis *Processos* e *Posição relativa* correlacionam-se à CN explícita: plural regular e itens pospostos ao núcleo desfavorecem a aplicação do processo.

A comparação dos três modelos de regressão logística de efeitos mistos mostra que, de fato, as correlações das variáveis consideradas na análise da CN explícita são afetadas pela presença de ambas as variáveis *Processos* e *Tonicidade* em um modelo e pela ausência de uma delas nos dois outros. Especialmente no que diz respeito às variáveis previsoras sociais, essas exibem correlação com a CN explícita apenas nos modelos em que *Tonicidade* está presente (tabelas 5 e 6).

No entanto, o mais importante a observar na comparação dos três modelos refere-se às estimativas: elas são coerentes. Não há alteração no valor das estimativas nos níveis das diferentes variáveis previsoras e naquilo que as estimativas apontam: valor negativo,

desfavorecimento da variável resposta; positivo, favorecimento da variável resposta. Isso permite enunciar generalizações sobre os resultados,<sup>14</sup> considerando o que revela o conjunto dos três modelos.

Nos modelos com Tonicidade (tabelas 5 e 6), a CN explícita é explicada como efeito da ação de variáveis sociais e linguísticas, especialmente Renda, Escolaridade, Posição Relativa, Processos e a própria Tonicidade. Já o modelo com Processos, sem Tonicidade (tabela 7) mostra que as variáveis Posição Relativa e Processos têm maior poder explicativo sobre a CN explícita do que as demais variáveis previsoras juntas. Esse achado, no entanto, não invalida o que revelam os modelos com Tonicidade (tabelas 5 e 6): essa variável tem efeito sobre a CN explícita, mas menor do que o de Processos, razão pela qual as demais variáveis controladas exibem correlação quando apenas Tonicidade está presente (tabela 6).

A análise quantitativa aqui realizada revela, então, que a CN explícita, bastante frequente no português de porto-alegrenses, é favorecida por elementos pospostos ao núcleo do sintagma nominal e por falantes com nível superior de escolaridade, é desfavorecida por plural regular, vocábulos paroxítonos e proparoxítonos e falantes de renda baixa. Sobre as variáveis Gênero e Zona, que apresentaram correlação com a CN explícita apenas em um dos três modelos, vale dizer, considerando-se as proporções de aplicação na amostra, que há mais marcação de plural na fala de mulheres do que na de homens, e na de falantes que residem no Centro de Porto Alegre do que nas demais zonas.

Esses resultados não apenas confirmam nossas hipóteses de pesquisa, mas também se equiparam ao atestado em outros estudos sobre a variação de número no PB – Scherre (1988), Oushiro (2015), Mangabeira (2016), citando apenas alguns. Conformam-se à constatação de Scherre (1994) sobre a CN variável como processo de variação inerente no PB: manifesta-se em contextos linguísticos e sociais similares, com tendências a CN explícita ou zero bastante previsíveis.

Em relação às variáveis previsoras sociais, especialmente *Escolaridade*, vale dizer que esperávamos o papel favorecedor do maior nível de escolarização sobre a CN explícita, pela maior exposição dos sujeitos à cultura letrada e maior realização de práticas mediadas pela leitura e pela escrita, tanto ao longo da formação escolar quanto nas atividades socioeconômicas realizadas a partir dela. Como Scherre (1998a) e Scherre e Naro (2006), nossos resultados mostram que as proporções de CN explícita aumentam conforme aumenta o nível de escolaridade. No que se refere à zona Centro, seus efeitos

---

<sup>14</sup> Agradecemos a Adalberto Ayjara Dornelles Filho (IBGE-Unidade Estadual/Rio Grande do Sul) a consultoria estatística na interpretação dos resultados dos três modelos.



favorecedores sobre a CN explícita comprovam que, em grandes centros urbanos brasileiros como Porto Alegre, as áreas centrais são o *locus* de práticas socioeconômicas relativamente privilegiadas, nas quais a CN explícita pode ser valorizada como símbolo de conhecimento (da linguagem padrão) e de autoridade (Bourdieu, 1996). Essas razões, referentes ao espaço físico, estendem-se também à renda, que financia o acesso ao espaço físico e explica o favorecimento da CN explícita pelo nível mais alto de renda, aferida, em nosso estudo, a partir da renda domiciliar média mensal das famílias no bairro onde reside o informante.

Os modelos estatísticos mostram, portanto, que a CN explícita é condicionada por variáveis sociais e linguísticas no português de Porto Alegre, de forma coerente com os estudos referidos até aqui. Nosso próximo passo é aprofundar, em alguma medida, o esclarecimento das motivações sociais do processo. Realizamos uma análise das práticas sociais de alguns dos informantes de que extraímos os dados para a análise quantitativa. Levamos em conta o declarado pelos informantes sobre suas práticas cotidianas nas entrevistas sociolinguísticas do LínguaPOA (2015-2019). Categorias correlacionadas à CN explícita, quer por favorecê-la, quer por inibi-la, serão mobilizadas na análise, mas agora em uma perspectiva distinta, qualitativa.

### **“Bares estilizados, bares mais bacaninhas”: os resultados da análise qualitativa**

A análise qualitativa sobre práticas sociais dos informantes em que a CN explícita ou zero realizam-se fundamenta-se na teoria social de Bourdieu (1996, 2015).

Uma noção fundamental dessa teoria é a de *habitus*, conjunto de disposições a agir socialmente que se desenvolve ao longo da vida de um indivíduo e está no centro da reprodução social. O *habitus* é um “gerador de práticas objetivamente classificáveis e, ao mesmo tempo, sistema de classificações [...] de tais práticas” (BOURDIEU, 2015, p. 162). É na relação entre essas duas dimensões, da prática e da percepção da prática, que se constituem os diferentes estilos de vida, definidos pelo sociólogo como resultados sistêmicos do *habitus*. Ao serem percebidos, os estilos de vida são qualificados socialmente.

No cerne dos estilos de vida, encontra-se o gosto, produto da diferenciação e apreciação das práticas. Caracterizado como operador prático, motivador da conversão das coisas em sinais distintivos e distintos, o gosto transforma práticas classificadas em manifestações simbólicas. Assim, concebem-se estilos de vida como “um conjunto unitário de preferências distintivas que exprimem, na lógica específica de cada um dos subespaços simbólicos – mobiliário, vestuário, linguagem e/ou hexis corporal – a mesma intenção

expressiva” (Bourdieu, 2015, p. 165). Ou seja, pode-se entender que variantes como a CN explícita integram estilos linguísticos concernentes a certos estilos de vida. São, assim, apreciadas ou, eventualmente, desvalorizadas, conforme o gosto dos estilos de vida particulares.

É desse modo que as trocas linguísticas, segundo Bourdieu (1996), vêm a ser trocas econômicas em mercados linguísticos. Localizadas em relações de força simbólica, as trocas linguísticas podem conferir valor material/simbólico ao discurso, a depender da posição e dos papéis dos sujeitos no espaço social, decorrentes da soma de seus capitais, especialmente do econômico (riquezas) e cultural (conhecimento, escolaridade). Assim, para Bourdieu, os discursos são signos de riqueza e de autoridade, implicam obediência e conferem confiabilidade. Os agentes sociais posicionados no topo da hierarquia social possuem um *habitus* linguístico resultante da antecipação dos lucros no mercado linguístico, uma vez que já conhecem a dinâmica, as leis e os valores das formas linguísticas que ali circulam.

Na teoria social de Bourdieu, portanto, um agente social – e, supostamente, seu padrão de fala – é percebido e classificado socialmente a partir da sua posição relativa no espaço social, dependente do tipo e quantia do capital que possui, a qual lhe confere certo estilo de vida e lhe fornece o gosto, um esquema de apreciação/ação. As classes dominantes são marcadas pelo que Bourdieu (2015) denomina gosto de luxo, as classes dominadas, por estilos de vida calcados no gosto de necessidade. Cabe-nos questionar quais são os estilos de vida dos 8 informantes, do total de 32, cujas entrevistas sociolinguísticas são examinadas qualitativamente aqui, por análise de conteúdo. Objetivamos esclarecer, a esse respeito, as práticas sociais em que as variantes da CN, explícita ou zero, tendem a manifestar-se.

Oliveira (2018), sobre variação linguística e significados sociais no PB de POA, identificou, entre os informantes do LínguaPOA investigados, pelo menos dois estilos de vida:

grupo A [...], composto por pessoas que, dentre outros aspectos, circulam a pé pelo centro da cidade (onde realizam práticas culturais), não ouvem rádio, tendem a ser favoráveis à legalização da maconha e mencionam políticos de partidos de esquerda como bons exemplos; grupo B [...], composto por pessoas que, dentre outros aspectos, circulam de carro na cidade (não no centro, mas preferencialmente na orla, onde praticam esporte ao ar livre, como corrida), ouvem rádio, são desfavoráveis à legalização das drogas e têm dificuldade em mencionar bons exemplos de políticos. (grifos nossos) (Oliveira, 2018, p.8)

Os informantes cujas entrevistas sociolinguísticas examinamos na análise de conteúdo, desta vez discutindo também a CN zero, não são exatamente os mesmos considerados por Oliveira (2018). Por essa razão, não operamos com os estilos de vida A e B identificados pelo autor. Mas chamam atenção, no contraste dos dois estilos, as diferenças identitário-ideológicas e as práticas sociais distintas realizadas pelos sujeitos, que podem dar origem a significados sociais peculiares, associados a uma mesma forma linguística, como a CN zero ou a CN explícita. Chama atenção, também, o fato de um dos grupos realizar práticas culturais no centro da cidade, outro, não.

Como afirmamos anteriormente, a CN explícita, requerida e mantida por práticas letradas fomentadas pela escola, pode ser socialmente prestigiada em certos estilos de vida, desprestigiada em outros. Práticas culturais que requerem o uso da CN explícita dependem de sua vinculação mais ou menos estreita com a escolarização de quem as realiza.

O incremento da escolarização da população vem sendo apontado na literatura como crucial para o aumento da CN explícita entre os brasileiros: “a variação na concordância nominal de número [...] não reflete mudança clara para todos os falantes [...], embora estejamos capturando aumento de concordância em função de maior exposição ao ambiente escolar”. (Scherre; Naro, 2006, p. 120). Na escola, ensina-se a norma-padrão da Língua Portuguesa, o que, segundo Dália e Lucchesi (2020), pode ter efeito nas práticas linguísticas dos sujeitos: “[...] a variação apresentada pela geração de jovens, na atualidade, pode estar associada ao seu maior acesso à educação formal [...] fomentando, na maioria das vezes, o estigma social sobre as formas da linguagem popular e o preconceito linguístico.” (Dália; Lucchesi, 2020, p. 232).

No entanto, gostaríamos de acrescentar outro fator a essa discussão: acreditamos que, para além da exposição à educação formal como fator de promoção da CN explícita, a renda, a localização dos sujeitos no espaço urbano (moradia) e os espaços sociais e institucionais frequentados por eles também interferem na variável, na medida em que esses fatores definem seus estilos de vida. Sabemos que a convivência nesses espaços, na maioria das vezes, depende do nível de escolaridade e da renda. Em outras palavras, o acesso à educação formal, especialmente em níveis superiores, abre portas no mercado de trabalho, viabilizando salários mais altos e, conseqüentemente, estilos de vida relativamente privilegiados, pelo aumento de capital (cultural e econômico, principalmente). Isso porque, esperada na linguagem padrão, a CN explícita indexa autoridade e confere prestígio aos falantes que a utilizam em locais onde tal uso é esperado.

Como explicamos na seção Metodologia, os informantes selecionados para a aná-

lise qualitativa apresentam proporções distintas de CN zero. Distribuem-se ao longo de um *continuum*, que vai de 0,7% a 44% de CN zero (Tabela 1). Seleccionamos focalizar a discussão nos dois informantes das pontas (Inf09 e Inf62) e nos dois do meio (Inf45 e Inf26) do *continuum*, isto é, em extremos opostos e na posição intermediária do *continuum*, para contrastar as diferenças e semelhanças na aplicação de CN zero, buscando esclarecer tais contrastes por associação a práticas sociais dos informantes.

Iniciaremos pelos informantes 09 (escolaridade Superior, bairro renda média alta, zona Centro, 0,7% de CN zero) e 62 (escolaridade Básica, bairro renda média baixa, zona Norte, 44% de CN zero), nos extremos do *continuum* de CN zero. Eles têm também perfis sociais contrastantes: de um lado está o informante 09, formado em Direito e servidor público do estado do Rio Grande do Sul, com um uso de CN zero abaixo de 1%; de outro lado, o informante 62, trabalhador autônomo na área de manutenção elétrica, com um uso de CN zero acima de 40%. Quanto ao espaço geográfico (figura 1), podemos observar a distância do bairro Cristo Redentor, onde reside o informante 62, em relação ao centro, e sua proximidade com os limites ou periferia da cidade. Esses fatores corroboram nossa hipótese, de que escolaridade e renda, assim como a relação centro-periferia, condicionam a CN variável.

A análise de conteúdo revela que o informante 09 realiza diversas práticas sociais de um estilo de vida pautado pelo gosto de luxo (Bourdieu, 2015). Ele relata ter duas bandas e a oportunidade de tocar em espaços públicos, como bares. Além disso, apesar de ser formado em Direito e exercer a profissão, seu grupo social de amigos é formado majoritariamente por músicos, com quem frequenta muitos bares na Cidade Baixa, bairro boêmio da cidade. Além disso, gosta de frequentar, com a esposa, locais como restaurantes, cafeterias e cinemas. Por fim, o informante declara que gosta muito de viajar, já visitou muitos estados brasileiros, assim como diversos países europeus e os Estados Unidos. Já o informante 62, que apresenta escolaridade básica, renda baixa e reside longe do centro da cidade, realiza práticas sociais de um estilo de vida orientado pelo gosto de necessidade (Bourdieu, 2015). Ele considera-se uma pessoa “caseira”. Declara a todo o momento, ao longo da entrevista, que não gosta de sair de casa, não gosta de viajar, não frequenta espaços públicos. Resumidamente, o informante afirma ser uma “pessoa que não se mistura”.

A análise de conteúdo das entrevistas dos informantes 09 e 62, portanto, vai ao encontro de nossa hipótese: o conjunto das diferenças de escolaridade, renda e local de moradia associa-se a práticas sociais distintas, peculiares a estilos de vida e gostos diversos, incluindo-se nessas práticas a fala e as formas linguísticas habitualmente empregadas,

como a CN explícita ou zero. No entanto, os achados da análise das fichas sociais e do conteúdo das entrevistas dos informantes 45 e 26, de certa forma, resistem parcialmente à nossa tese, já que, embora seus perfis e práticas sejam distintos, apresentam proporções de CN zero praticamente iguais.

Os informantes 45 (escolaridade Superior, bairro renda média alta, zona norte, 13% de CN zero) e 26 (escolaridade Básica, bairro renda média baixa, zona central, 14% de CN zero) apresentam diferenças quanto à escolaridade e à renda. O informante 45 trabalha em um centro de pesquisa da PUCRS, enquanto o informante 26 é empresário e sócio de uma rede de lancherias em Porto Alegre. Se os perfis sociais são diferentes, a que se devem as proporções similares de CN zero?

O informante 26, embora tenha escolaridade básica, tem uma ocupação profissional (empresário) detentora de capital (econômico). Além disso, o bairro onde mora, mesmo que seja de renda média baixa, situa-se no centro da cidade. Assim, costuma frequentar *shoppings*, viajar (especialmente para ver o filho), ir a bares e restaurantes com os amigos. Já o informante 45, que tem maior escolaridade (Superior) e ocupação identificada com a elite (trabalha em um centro de pesquisa universitário), afirma não gostar de viajar, nem de sair. Raramente vai a restaurantes com a família. Declara-se uma pessoa caseira, com interesses em política e economia, mesmo que não se envolva em grupos de discussão acerca desses temas. Ou seja, os informantes 26 e 45 mesclam práticas dos estilos de vida pautados pelos gostos de luxo e de necessidade, o que pode explicar as proporções similares de CN zero, mesmo tendo traços de perfil social distintos.

Os achados da análise de conteúdo dessas entrevistas contribuem para esclarecer de que forma variáveis como Escolaridade e Renda, correlacionadas ao uso variável da CN, operam conjuntamente na CN variável: elas estão na base das disposições do *habitus*. Definem o ponto, no espaço social, ocupado pelos sujeitos, seu local de moradia, estilos de vida e gostos. São compatíveis com a relação centro-periferia estabelecida, pelos informantes, em suas práticas sociais na cidade. Quanto mais um sujeito participa de espaços sociais que privilegiam a linguagem padrão, mais ele usa a CN explícita. Portanto, como afirmamos anteriormente, acreditamos que, além da escolaridade, fatores como renda, zona (ou local de moradia) atuam conjuntamente na CN variável no PB, por incorporarem-se a estilos de vida e gostos em cujos padrões de fala as variantes CN explícita e CN zero podem ser socialmente significativas.

### **Últimas palavras**

Os resultados da análise da CN variável no PB de POA, realizada em uma perspecti-

va mórfica, mostraram que a proporção de CN explícita é alta no português dessa comunidade de fala. A variação observada é condicionada por variáveis linguísticas e sociais. Nos diferentes modelos estatísticos testados, verificou-se que a CN explícita é favorecida, no PB de POA, pela posição posposta ao núcleo do sintagma e por falantes com ensino superior de escolaridade. É desfavorecida por plural regular, paroxítona e proparoxítona, falantes de renda baixa. Os resultados da análise quantitativa de produção são coerentes, portanto, com os resultados de estudos da CN variável em outras comunidades de fala brasileiras, indo ao encontro da afirmação de Scherre (1994) de que o processo apresenta tendências sistemáticas e previsíveis.

A análise de conteúdo lançou luz à associação entre a CN variável, escolaridade, renda e local de moradia dos informantes, em termos de práticas sociais e estilos de vida. A correlação entre maior escolaridade e maior frequência de CN explícita, atestada na análise quantitativa, nem sempre se comprova no exame de dados de informantes individuais. Os relatos nas entrevistas sociolinguísticas mostram que, se a maior escolaridade leva à maior CN explícita, esta é atributo de informantes cujos estilos de vida pautam-se pelo que se poderia denominar, com Bourdieu (2015), de gosto de luxo, levando-os a circular pelo centro da cidade e consumir produtos culturais, em práticas que valorizam e requerem o uso da CN explícita.

Para ampliar a compreensão da CN variável no PB de POA, teria sido necessário incluir, no artigo, também os resultados da análise de produção na perspectiva *sintagmática*, que, como afirmamos, não se fez pela limitação de espaço no artigo, mas encontra-se em Santos (2022). Faltou ao artigo, além disso, controlar outros grupos etários que não apenas o de informantes com 40 a 59 anos de idade, como fizemos aqui. Isso teria oportunizado investigar a CN variável na perspectiva da mudança em tempo aparente, com dados de entrevistas disponíveis no mesmo banco de dados, o LínguaPOA (2015-2019). Esta etapa fica como tarefa de pesquisa futura e tema de publicações que, esperamos, venham adiante.

## Referências

- BARDIN, L. *Análise de conteúdo*. Ed. revista e ampliada. Trad. Luis Antero Reto, Augusto Pinheiro. São Paulo: Edições 70, 2011.
- BATTISTI, E.; OUSHIRO, L. A motivação social da haplologia varia no português de Porto Alegre. *Confluência*, n. 62, p. 270-302, 2022.



BORTONI-RICARDO, S. M. *Nós cheguei na escola, e agora?* São Paulo: Parábola Editorial, 2005.

BOURDIEU, P. *A economia das trocas linguísticas*. São Paulo: Editora da Universidade de São Paulo, 1996.

BOURDIEU, P. *A distinção: crítica social do julgamento*. Trad. Daniela Kern, Guilherme J. F. Teixeira 2. ed. rev. 2. reimpr. Porto Alegre: Zouk, 2015.

CAMARA JR., J. M. *Estrutura da língua portuguesa*. 34. ed. Petrópolis: Vozes, 2015.

CAVALCANTI, J. As faces de uma polêmica: o episódio do livro didático “Por uma vida melhor”. *D.E.L.T.A*, v. 29, Edição Especial, p. 487-501, 2013.

CEDERGREN, H.; SANKOFF, D. Variable rules: performance as a statistical reflection of competence. *Language*, v. 50, n. 2, p. 333–355, 1974.

CHAVES, R. Q. Princípios de Saliência Fônica: isso não soa bem. *Letrônica*, v. 7, n. 2, p. 522-550, jul./dez., 2014.

DÁLIA, J; LUCCHESI, D. A variação na concordância de número do sintagma nominal no português rural da serra fluminense: deriva ou contato? *Gragoatá*, v. 26, n. 54, p. 217-251, 2021.

LABOV, W. *Padrões sociolinguísticos*. Trad. Marcos Bagno, Maria Marta Pereira Scherre e Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008.

LABOV, W. *Principles of linguistic change: internal factors*. Cambridge/Oxford: Blackwell, 1994.

LABOV, W. *Principles of linguistic change: social factors*. Malden/Oxford: Blackwell, 2001.

LABOV, W. *Principles of linguistic change: cognitive and cultural factors*. Malden/Oxford/West Sussex: Wiley-Blackwell, 2010.

LÍNGUAPOA. Universidade Federal do Rio Grande do Sul. 2015-2019 (período de coleta). Disponível em: <https://www.ufrgs.br/linguapoa/>. Acesso em: 16/06/2022.

LUCCHESI, D. Ciência ou dogma? O caso do livro do MEC e o ensino de Língua Portuguesa no Brasil. *Revista Letras*, n. 83, p. 163-187, 2011.

LUCCHESI, D; DÁLIA, J. Novos condicionamentos estruturais da variação na concordância nominal de número. *Forum Linguístico*, v. 19, n. 1, p. 7369-7386, 2022.

LUFT, C. P. *Moderna Gramática Brasileira*. 2. Ed. revisada e atual. São Paulo: Globo, 2002.

MANGABEIRA, A. B. de A. *Variação na concordância nominal, prática social e identidade entre jovens e adultos do Centro do Trabalhador (Porto Alegre - RS)*. 2016. 319 f. Tese (Doutorado em Linguística Aplicada) - Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2016.

NARO, A. J.; SCHERRE, M. M. P. Sobre as origens do português popular do Brasil. In: NARO, A. J.; SCHERRE, M. M. P. (orgs.). *Origens do português brasileiro*. São Paulo: Parábola, 2007. p. 25-48.

OLIVEIRA, S. G. de. *Ingliding de vogais tônicas como prática estilística no falar porto-alegrense: significados sociais da variação linguística*. 2018. 230 f. Dissertação (Mestrado em Letras) – Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2018.

OUSHIRO, L. *Identidade na Pluralidade: avaliação, produção e percepção linguística na cidade de São Paulo*. 2015. 349 f. Tese (Doutorado em Linguística) - Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2015.

PONTES, V. M. *A concordância nominal de uma comunidade de Porto Alegre*. 1979. Dissertação (Mestrado em Letras) - Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1979.

SANKOFF, D.; TAGLIAMONTE, S.; SMITH, E. *GoldVarb X: Variable Rule Application for Macintosh and Windows*. Toronto: University of Toronto, 2005.

SANTOS, B. S. dos. *Os mesmos desfavorecido: a variação na concordância nominal de número em Porto Alegre/RS*. Trabalho de Conclusão de Curso (Licenciatura em Letras) – Instituto de Letras, Universidade Federal do Rio Grande do Sul, 2022.

SCHERRE, M. M. P. *Reanálise da concordância nominal em português*. 1988. 555 f. Tese (Doutorado em Linguística) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1988.

SCHERRE, M. M. P. Sobre a atuação do princípio da Saliência Fônica na concordância nominal. In: TARALLO, F. (org). *Fotografias sociolinguísticas*. Campinas, SP: Pontes, 1989. p.301-302.

SCHERRE, M. M. P. Aspectos da concordância de número no português do Brasil. *Revista Internacional de Língua Portuguesa - Norma e Variação do Português*, p. 12-37-49, 1994.

SCHERRE, M. M. P. Sobre a influência de variáveis sociais na concordância nominal. In: SILVA, G. M. de O; SCHERRE, M. M. P. *Padrões sociolinguísticos: análise de fenômenos variáveis no português falado na cidade do Rio de Janeiro*. Rio de Janeiro: Tempo Brasileiro, RJ, 1998a. p.239-264.

SCHERRE, M. M. P. Sobre a influência de três variáveis relacionadas na concordância nominal em português. In: SILVA, G. M. de O; SCHERRE, M. M. P. *Padrões sociolinguísticos: análise de fenômenos variáveis no português falado na cidade do Rio de Janeiro*. Rio de Janeiro: Tempo Brasileiro, RJ, 1998b. p. 85-117.

SCHERRE, M. M. P.; NARO, A. Sobre a concordância nominal de número no português falado no Brasil. In: RUFFINO, G. (org). *Dialetologia, geolinguística, sociolinguística*. (Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza) Centro di Studi Filologici e Linguistici Siciliani, Università di Palermo. Tübingen: Max Niemeyer Verlag, 5:509-523, 1998.

SCHERRE, M. M. P.; NARO, A. Mudança sem mudança: a concordância nominal de número no português brasileiro. *Scripta*, v. 9, n. 18, p. 107-129, 2006.

THE R CORE TEAM. *R: a language and environment for statistical computing*. Viena: R Foundation for Statistical Computing, 2019. Disponível em: <https://www.R-project.org>. Acesso em: 16/06/2022.

