



# Análise contrastiva da estrutura do sintagma nominal possessivizado no português brasileiro

Manoel Siqueira

Universidade Federal de Sergipe

<https://orcid.org/0000-0002-5928-3450>

## RESUMO

Neste trabalho, comparamos resultados de três conjuntos de dados sobre a estrutura de sintagmas nominais possessivizados no português brasileiro falado, com especial atenção para os possessivos antepostos a nomes, como em *sua casa*, e possessivos pospostos a nomes, como em *uma casa sua*, de modo a observar se é possível apresentar uma generalização para o comportamento dessa estrutura na língua. Usamos como base para a comparação as descrições feitas por Neves (1993) e Franchi (1996), que trabalham com dados de fala do NURC, e dados de fala do banco *Falares Sergipanos*. Como método, adotamos o descritivo/inferencial, de modo a observar a distribuição e a associação entre os dados. Os testes de associação apontam que, em ambos os tipos de possessivos, há mais diferenças entre as três amostras e seus resultados do que semelhanças, o que não nos permite desenvolver uma generalização do comportamento desses SN para o português.

**PALAVRAS-CHAVE:** Análise contrastiva; SN possessivizado; Português brasileiro.

## Contrastive Analysis of The Structure of The Possessivized Noun Phrase in Brazilian Portuguese

### ABSTRACT

In this paper, we compared the results of three data sets about the structure of possessive noun phrases in spoken Brazilian Portuguese, with special attention to pre-nominal possessives, as in *sua casa*, and post-nominal possessives, as in *uma casa sua*, in order to observe if it is possible to present a generalization for the behavior of this structure in the language. We used as a basis for comparison the descriptions made by Neves (1993) and Franchi (1996), who worked with speech data from NURC, and speech data from *Falares Sergipanos* data. The methodology adopted was the descriptive/inferential analysis, in order to observe the distribution and the association among the data. The association tests show that, in both types of possessives, there are more differences between the three samples and their results than similarities, which does not allow us to develop a generalization about the behavior of these NP to Portuguese.

**KEYWORDS:** Contrastive analysis; Possessivized NP; Brazilian Portuguese.



## 1. Introdução

No português brasileiro, sintagmas nominais possessivizados – que possuem no núcleo nominal um possessivo (*meu, seu, nosso* etc.) – podem ser estruturados como Possessivo + Nome, em que o possessivo antecede o nome, como em (1), e como Nome + Possessivo, em que o possessivo procede o nome, como em (2).

(1) mesma casa e eu ia aguentar até *meu*<sub>Poss</sub> *nome*<sub>N</sub> sair na lista. (ADR1MI<sup>1</sup>)

(2) entre eu e um *colega*<sub>N</sub> *meu*<sub>Poss</sub> que é capitão. (CAS1MF)

Dentro dessas estruturas, o sintagma nominal possessivizado (SN possessivizado) pode ser organizado contendo artigos definidos (*a, o, as, os*), artigos indefinidos (*um, uma, uns, umas*), demonstrativos (*esse, este, aquele* etc.), numerais (*dois, três, quatro* etc.), o quantificador “todo” (e variantes), entre outros elementos. Estudos que descrevem a organização dos tipos de SN possessivizados no português, como os de Neves (1993) e Franchi (1996), observam que há uma certa rigidez na organização: (i) em construções do tipo Possessivo + Nome predomina o uso de artigo definido; (ii) em construções do tipo Nome + Possessivo predomina o uso de não-definidores.

Contudo, a ausência de uma agenda unificada em relação à pesquisa sociolinguística no Brasil dificulta a comparabilidade e a generalização de resultados para o comportamento do português brasileiro (FREITAG e ROST-SNICHELOTTO, 2015), visto que diferenças/igualdades podem ser condicionadas tanto pela metodologia da amostra quanto por fatores internos e externos à língua. Neves (1993) e Franchi (1996), por exemplo, usam dados do Projeto Norma Urbana Culta (NURC), constituído pela fala de “informantes cultos”, com ensino superior completo, focando em elocuções formais, organizado na década de 70. Um protocolo similar ao do NURC ajudaria no desenvolvimento de comparações.

Neste trabalho, desenvolvemos uma análise contrastiva entre descrições sobre a organização de sintagmas nominais possessivizados no português brasileiro. Utilizamos dados de trabalhos feitos com base no NURC, como Neves (1993), Franchi (1996) e Müller (1998), juntamente a dados retirados do banco Falares Sergipanos (FREITAG, 2013; 2017), que também considera a modalidade de língua dita “cultu”: utiliza-se a fala de estudantes do ensino superior e falantes com superior completo. Indagamos se a rigidez na organização desses sintagmas é uma particularidade do português brasileiro ou apenas da amostra do NURC. Nossa hipótese é a de que as diferenças observadas na organização dos SN possessivizados são resultante da antiguidade da amostra do NURC, considerando que as duas amostras aqui apresentadas foram constituídas em tempos diferentes, e que a língua muda no tempo e espaço (LABOV, 1994).

<sup>1</sup> Os exemplos citados neste trabalho que possuem esse tipo de codificação entre parênteses são retirados das entrevistas sociolinguísticas que compõem o *corpus* desta pesquisa, pertencentes ao Banco de Dados Falares Sergipanos (FREITAG, 2013; 2017).

## 2. Ordem dos elementos do SN possessivizado

Em português, os possessivos podem ocorrer em i) posição pré-nominal, como em (3); ii) em posição pós-nominal, como em (4); iii) suprimindo um nome já estabelecido no discurso, como em (5); iv) em contextos predicativos, como em (6)<sup>2</sup>.

(3) Meu livro.

(4) Um livro meu.

(5) Teu livro e o meu (livro).

(6) Este é o meu livro/Esse livro é o meu (livro).

Em (3), o possessivo “meu” antecede o nome “livro”, um possessivo pré-nominal ou anteposto. Em (4), há uma permutação, em que o possessivo “meu” aparece depois do nome, um possessivo pós-nominal ou posposto. Em (5), o possessivo “meu” com determinante definido o permite a elipse do nome no segundo possessivo, uma vez que o nome “livro” se encontra presente na projeção discursiva. Por último, em (6), o possessivo “meu” encontra-se na posição predicativa, havendo uma ligação entre o possessivo e o sintagma nominal (SN), já presente na sentença. Embora haja essas posições, nosso enfoque recai sobre as posições (3) e (4), os possessivos antepostos e pospostos ao nome, frente a nossa indagação inicial sobre a organização dos SN com essas posições, e para manter linearidade com as demais pesquisas sobre a temática (FRANCHI, 1996; NEVES, 1993).

Segundo Perini (2005, p. 94), “o SN tem uma estrutura posicionalmente muito mais rígida do que a oração; as possibilidades de mudança de ordem dos termos são poucas e bem delimitadas”. Para o autor, não são muitas as classes de elementos que podem acompanhar um nome num sintagma. Essas classes são limitadas e a movimentação desses elementos restringe-se a algumas poucas posições, quando não a apenas uma posição.

Perini (2005) propõe a ideia de SN máximo para demonstrar as posições das classes de palavras que ocorrem num sintagma nominal. Um SN máximo é uma abstração linguística que representa um sintagma nominal em que todas as posições estão preenchidas. Conta-se, numa sequência linear, a posição dos elementos com base na distância entre esses elementos e o nome núcleo do SN máximo, cada elemento representa uma posição. Observemos os seguintes exemplos:

(7) Meu carro.

(8) Aquele meu carro.

No exemplo (7), o possessivo “meu” ocorre logo à primeira posição à esquerda do nome “carro”. Contudo, em (8), “meu” está ocupando a segunda posição, já que o demonstrativo “aquele” ocupa a primeira posição: “aqui já temos um problema: parece que a posição de ‘meu’ é variável” (PERINI, 2005, p. 95). O autor complementa a ideia ao dizer que “uma observação que nos pode ajudar é

<sup>2</sup> Há outros contextos de ocorrências, como em vocativos (*meu* filho, *venha cá!*) – e expressões cristalizadas (*Meu* Deus!). Apresentamos somente esses contextos, pois são os que apresentam variação em suas realizações.

a seguinte: ‘meu’ pode ocorrer em primeiro lugar, mas nunca quando o sintagma também tem o item ‘aquele’” (PERINI, 2005, p. 95). Construções como as em (9) são agramaticais, i.e., não pertencentes à estrutura da língua.

(9) \*Meu aquele carro.

Chegamos à seguinte observação: i) um demonstrativo, como “aquele”, sempre ocupa a primeira posição, enquanto um possessivo, como “meu”, ocupa sempre a segunda (PERINI, 2005). Esses elementos não precisam necessariamente estar presentes para sabermos sua posição: “é como se considerássemos que existe uma posição ‘não-preenchida’ antes de ‘meu’” (PERINI, 2005, p. 95). Há sintagmas em que meu pode ser o primeiro elemento, como (7), como também há sintagmas em que “meu” pode estar em segundo, como (8). Para Perini (2005, p. 95), “trata-se, portanto, de uma espécie de fila de preferências para ocupar a primeira posição”. Agora observemos os exemplos abaixo:

(10) Meus amigos.

(11) Aqueles meus amigos

(12) Todos meus amigos

(13) Os meus amigos.

(14) Meus dois amigos.

(15) ?Um meu amigo.

No exemplo em (10), vemos o contexto já apresentado antes, em que o possessivo “meus” ocorre na primeira posição. Vimos também no exemplo (11), que a primeira posição é ocupada por “aquele” e o possessivo “meus” encontra-se na segunda. Em (12), a primeira posição é ocupada por um quantificador, “todos”, enquanto “meus” encontra-se na segunda. A posição do possessivo também se repete em (13), em que a primeira posição é preenchida pelo determinante definido “o”. Em (14), o possessivo preenche a primeira posição, a segunda posição é aqui preenchida pelo numeral “dois”. Por fim, em (15), a primeira posição é preenchida pelo determinante indefinido “um”, e o possessivo preenche a segunda posição. O símbolo (?) indica dúvida em relação a construção, se agramatical ou não. Esses exemplos demonstram que a posição do possessivo em relação ao nome é variável, mas que, ainda assim, há uma distribuição regular para essas posições. Os exemplos acima representam os possessivos antepostos.

Neves (2008) apresenta a seguinte distribuição para os elementos que podem ocorrer juntamente ao possessivo anteposto ao nome:

**QUADRO 1.** Ordem dos elementos do SN possessivizado com o possessivo anteposto ao substantivo

1	2	3	4	5
todo/ø	art. definido/art. indefinido/ demonstrativo/ø	POSSESSIVO	numeral ordinal/quantificador indefinido/outro/próprio/ø	NOME

Fonte: NEVES (2008, p. 581)



A ordem apresentada em Neves (2008) não determina que obrigatoriamente deva ocorrer apenas um desses elementos, ou todos eles. Em construções da língua, pode haver a combinação de mais de um desses elementos, como também só ocorrer a construção Poss + N (cf. 19):

- (16) Todos os meus livros.
- (17) Os meus dois livros.
- (18) Todos esses meus dois livros.
- (19) Meus livros.

Em (16), além do possessivo, há também o quantificador “todos” e o determinante definido “os”. A sentença em (17) possui, junto ao possessivo, o determinante definido “os” e o numeral “dois”. A sentença em (18) apresenta quatro elementos antes do nome: quantificador (todos), demonstrativo (esses), possessivo (meus) e numeral (dois). Por sua vez, (19) possui apenas o possessivo e o nome. Com isso, as opções de preenchimento no sintagma nominal possessivizado são variáveis para o possessivo anteposto.

Para o possessivo posposto ao nome, a organização do SN possessivizado é diferente do que vemos no Quadro 1. Vejamos os exemplos abaixo:

- (20) Um amigo meu trabalha lá.
- (21) O amigo meu trabalha lá.
- (22) Qualquer amigo meu trabalha lá.
- (23) Amigo meu não trabalha lá.
- (24) Dois amigos meus trabalham lá.

Como o possessivo é posposto ao SN, os elementos que alternam ocorrem antes do nome, enquanto o possessivo sempre ocorre após o nome. No exemplo em (20), a primeira posição é preenchida pelo determinante indefinido “um”. Em (21), a posição é preenchida pelo determinante definido “o”. Em (22) e (24), as primeiras posições são preenchidas por “qualquer” e o numeral “dois”. Já em (23), não há nenhum elemento preenchendo a primeira posição, ocorrendo apenas o nome.

Neves (1993) apresenta a seguinte organização para o SN possessivizado posposto ao nome (Quadro 2):

**QUADRO 2.** Ordem dos elementos do SN com o possessivo posposto ao substantivo

1	2	3	4
art. definido/art. Indefinido/ demonstrativo/ <i>qualquer</i> /ø	ø	SUBSTANTIVO	POSSESSIVO
ø	Numeral cardinal		

Fonte: NEVES (2008, p. 582)

Da mesma forma que no possessivo anteposto, a ordem dos elementos não implica que ocorra apenas um elemento por vez e a construção com apenas N + Poss é possível. Em exemplos

como (25), ocorre tanto o determinante definido “as” quanto o numeral “duas”. Já em exemplos como (26), temos o quantificador “todas” mais o determinante definido “as”. Em exemplos como (27), ocorre apenas N + Poss:

(25) As duas casas nossas foram vendidas.

(26) Todas as casas nossas foram vendidas.

(27) Casas nossas foram vendidas.

As opções de preenchimento no sintagma nominal possessivizado também são variáveis para o possessivo posposto. De forma geral, os dados observados em Neves (1993) apontam as possíveis realizações de elementos nos SN possessivizados. A ideia é que essa representação se adéque a toda a língua.

De fato, não esperamos que seja comum encontrar todos esses elementos. Formas como em (21), (25), (26) e (27) podem aparecer pouco na língua falada, como também possessivos pospostos ocorrem em menor número que os antepostos (NEVES, 1993; FRANCHI, 1996). As informações de Neves (1993; 2008) nos guiam para compreendermos a organização dos SN possessivizados, o que pode nos ajudar em nossa comparação. Contudo, precisamos considerar que essas informações foram coletadas há mais de 50 anos, com base no português ‘culto’ extraído do NURC. Línguas variam, ou seja, apresentam diferentes modos de dizer algo com o mesmo valor de verdade (LABOV, 1978) ao longo de sua história. E essa estrutura pode ter apresentado mudança desde a coleta de dados na década de 70. A mudança linguística faz parte de sua estrutura, o que pode indicar que o comportamento para a época da descrição apresentada por Neves (1993) não mais se aplique aos dias atuais, visto que a língua muda no tempo e no espaço (LABOV, 1994). É preciso, então, conhecermos as diferenças/semelhanças que podem estar presentes nos dados aqui utilizados.

### 3. Amostras sociolinguísticas no Brasil

A criação de banco de dados de língua falada no Brasil se desenvolveu por meio de dois aspectos metodológicos: o primeiro, seguindo procedimentos da Sociolinguística Variacionista, usa uma sistematização amostral de dados centrada num universo definido como comunidade de fala (FREITAG, 2016): grupos de indivíduos escolhidos não por compartilharem os mesmos traços linguísticos, mas sim por compartilharem os mesmos julgamentos sobre esses traços; o segundo usa uma técnica de amostragem denominada “estratificada aleatória” (BUCHSTALLER e KHATTAB, 2013; FREITAG, 2018): falantes são escolhidos aleatoriamente de modo a preencher células estratificadas segundo critérios sociodemográficos amplos, como sexo/gênero, escolaridade, faixa etária e zona de residência, com um número delimitado de participantes por célula. “Esse padrão de amostragem, por hipótese, confere confiabilidade e replicabilidade às análises” (FREITAG, 2018, p. 668), de modo que amostras realizadas em diferentes regiões poderiam dispor de uma forma mais confiável para se comparar resultados e desenvolver generalizações, frente à padronização na metodologia.



Contudo, no Brasil não houve (e ainda não há) uma padronização dos bancos de fala sociolinguísticos. A variação existente nos critérios para a constituição das amostras não permite a confiabilidade na comparação entre resultados de descrições. Diferenças/similaridades encontradas em análises e descrições podem ser resultantes tanto da metodologia amostral – sua organização, critérios sociodemográficos etc. – quanto de fatores internos e externos à língua: “só a padronização dos procedimentos metodológicos permitirá a realização de estudos contrastivos entre as variedades, para, então, possibilitar uma descrição mais acurada do português brasileiro” (FREITAG et al., 2012, p. 918), já que a metodologia da Sociolinguística Variacionista leva em conta a confiabilidade e a intersubjetividade (BAILEY e TILLERY, 2004).

É comum no desenvolvimento de estudos sociolinguísticos que outros estudos que abordem o mesmo tema, mas que analisam *corpora* distintos, sejam apresentados para dar suporte à discussão e à descrição dos dados por meio da comparação entre os resultados. Estudos seminais sobre a temática são candidatos a essa comparação e boa parte deles foi feita com base em dados de fala do projeto Norma Urbana Culta (NURC), que possibilitaram a descrição de diversos fenômenos da língua (cf. CASTILHO, 2012; JUBRAN, 2006; ILARI e NEVES, 2008; ILARI, 2014; 2015; NEVES, 2016). O projeto NURC conta com mais de 1500 horas de gravação realizadas na década de 70, com falantes de ambos sexos, nascidos nas capitais alvo (São Paulo, Rio de Janeiro, Porto Alegre, Recife, Salvador e Fortaleza), com ensino superior completo, distribuídos em três faixas etárias (25 a 35 anos; 36 a 55 e acima de 56). A fala desses informantes foi gravada durante aulas e conferências, como também diálogos informais e entrevistas. Objetivava-se obter dados que representassem o falar culto da época, a fala pertencente à parcela da população que tinha acesso aos mais altos níveis de ensino (CASTILHO e PRETI, 1986).

Embora o NURC não tenha seguido protocolos da metodologia sociolinguística para a coleta dos dados, as descrições feitas são basilares para revisões da literatura e revisões sistemáticas de variados fenômenos do português, já que “o acervo do Nurc permite que sejam realizadas – como de fato foram – diferentes abordagens para a descrição linguística, inclusive a sociolinguística de orientação variacionista” (FREITAG, 2019, p. 128).

Algumas amostras se aproximam da organização do NURC, como é o caso da amostra do projeto Falares Sergipanos (FREITAG, 2013; 2017), que considera uma modalidade de língua representante da visão de ‘falante culto’: a fala de estudantes universitários ou de pessoas com ensino superior completo. No entanto, os dados de fala são recolhidos por meio da “realização de entrevistas sociolinguísticas nos moldes labovianos, com estratificação sociodemográfica homogeneizada” (FREITAG, 2017, p. 122), o que difere da metodologia adotada no NURC (com dados de aulas, conferências, diálogos informais e elocuções formais), além de terem sido obtidos ao longo da última década.

Apesar disso, pesquisas desenvolvidas com base nesse *corpus* tendem a apresentar similaridades com resultados do NURC, como é o caso da pesquisa de Autor (ano) em que, apesar das diferenças metodológicas das amostras, os resultados quanto à distribuição dos possessivos de 3ª pessoa (seu/dele) se aproximam da distribuição em pesquisas do NURC (51,3% e 55,8% de “dele”, respectivamente), “o que refuta a hipótese “*corpus* velho” e reforça a hipótese de um conjunto de dados provenientes de entrevistas constituídas por temas mais referenciais” (FREITAG;

SIQUEIRA, 2018, p. 38). As similaridades dos resultados podem ser decorrente do tipo de metodologia escolhida para a constituição dos *corpora*.

Dessa forma, tendo em face a ideia da similaridade existente na constituição de ambos os *corpora*, podemos ter um melhor detalhamento contrastivo em relação à estrutura do sintagma nominal possessivizado no português brasileiro, o que pode nos dar indícios em relação ao seu comportamento no português.

#### 4. Conjunto de dados e método de análise

As diferenças/similaridades nos resultados de diferentes pesquisas podem ser observadas e descritas por meio de uma análise contrastiva entre os dados. Recorremos, neste trabalho, a uma meta-análise, procedimento no qual reunimos resultados de estudos com diferentes conjuntos de dados sobre a estrutura do SN possessivizado em português aplicamos técnicas estatísticas para explicar os resultados.

Usamos os dados obtidos por Neves (1993) e Franchi (1996) – amostra do NURC. Neves (1993) utiliza as seguintes gravações, retiradas do corpus mínimo do Projeto da Gramática do Português Falado: NURC: SP 234, 360, 405; BA 049, 098, 231; RS 045, 278, 291; RJ 328, 355, 379; PF 005, 131, 337; Franchi (1996), por sua vez, utiliza as seguintes gravações: BA 95; PA 6, 120, 291, 266, 365; PE 4, 79, 151, 266, 279, 340, 782; RJ 158, 328 c SP 62, 137, 161, 208, 234, 242, 250, 251, 255, 333, 343, 360, 396. Contrastamos os dados de Neves (1993) e Franchi (1996) com dados extraídos de 16 entrevistas sociolinguísticas retiradas da amostra Falares Sergipanos (FREITAG, 2013; 2017). Essas entrevistas foram submetidos ao programa LancsBox (BREZINA et al., 2020), *software* que realiza análises de língua em dados e em *corpora* e já faz a etiquetagem das palavras utilizadas na fala quanto à classe à qual pertencem, como a classe de possessivo.

Como havia sido feito em Neves (1993) e Franchi (1996), alguns contextos foram excluídos: contexto de vocativo, como “meu filho, vem cá”; expressões cristalizadas, como “Meu Deus do céu!”; contextos em que há a omissão do nome no segundo possessivo, como “trouxe meu livro e o teu”; e, por escolha nossa, contextos em que não podemos aferir se há ou não uso de determinantes, como “para ?a minha mãe, todos ?os meus parentes”, por considerarmos que pode ter ocorrido ou não uma assimilação.

Para o desenvolvimento da análise contrastiva, utilizamos técnicas estatísticas descritivas e inferenciais, para apresentar as medidas descritivas dos resultados, as frequências observadas, e para testar a associação entre os resultados. Realizamos teste de qui-quadrado, de modo a ver se a frequência absoluta observada de uma variável é significativamente diferente da distribuição da frequência absoluta esperada. Consideramos  $\alpha$  como 0,05<sup>3</sup>: um p-valor maior que 0,05 aponta que não há significância estatística na distribuição, enquanto um p-valor menor que 0,05 para significância. Usamos o  $V^2$  de Cramer para medir a associação entre as amostras. Essa medida

<sup>3</sup> O teste estatístico irá determinar um p-valor, que é comparado com um p-valor pré-determinado, conhecido como  $\alpha$  (alfa). Nas ciências humanas em geral, este valor é de 0,05. Isso significa que, se um teste for repetido 100 vezes, em cinco o resultado pode ser diferente do obtido inicialmente.



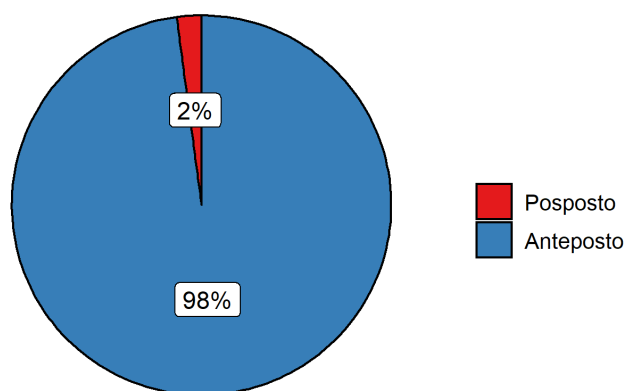
de associação varia de 0 a 1, em que: i) 0 é a ausência da associação; ii) valores altos do  $V^2$  de Cramer indicam uma relação mais forte entre as variáveis; iii) e os valores menores indicam uma relação fraca<sup>4</sup>. Os resultados são apresentados na seção 5.

## 5. Resultados e discussão

Após a etiquetagem, encontramos um total de 955 realizações de sintagmas nominais possessivizados nos dados do banco Falares Sergipanos (Gráfico 1):

**GRÁFICO 1.** Ocorrência de sintagmas nominais possessivizados no banco Falares Sergipanos

$$\chi^2_{\text{gof}}(1) = 876.68, p = 1.15\text{e-}192, \hat{V}_{\text{Cramer}} = 0.96, \text{CI}_{95\%} [0.78, 1.03], n_{\text{obs}} = 955$$



Fonte: Produzido pelo autor

A frequência de possessivos é menor com possessivos pospostos (2%  $n=20$ ), como em (28), do que com possessivos antepostos (98%  $n=935$ ), como em (29), e essa diferença é estatisticamente significativa, confirmada pelo valor de qui-quadrado, em que  $X^2(1, N=955) = 876.68$ ,  $p < 0,001$ , com associação forte ( $V^2 = 0.96$ ). Nos dados do português sergipano predomina o possessivo anteposto, na medida em que o possessivo posposto aparece em números baixos.

(28) essa visão tipo eu tenho “uma amiga minha” que ela fez Mecânica também. (CAS1MF)

(29) porque foi uma coisa que mudou “minha vida” que eu passei mais a ter. (ADR1MI)

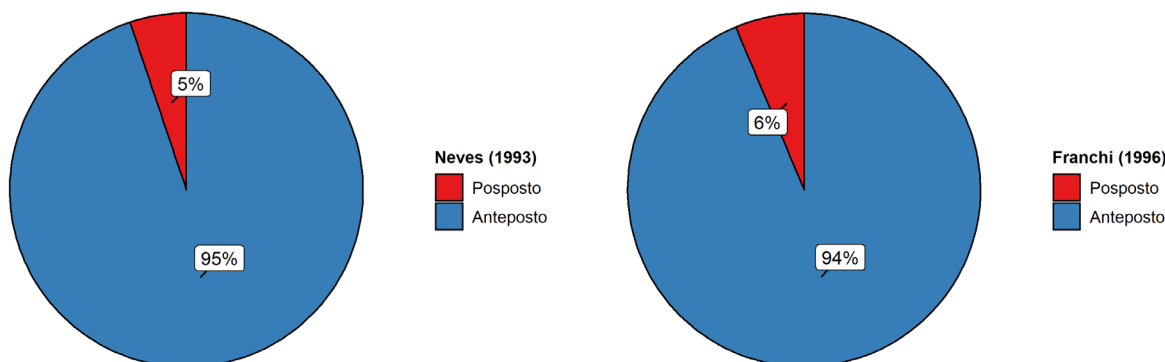
Observemos os números encontrados por Neves (1993) e Franchi (1996). Em Neves (1993), há 406 ocorrências de possessivos, enquanto em Franchi (1996) há um total 1136 ocorrências (Gráfico 2).

<sup>4</sup> Seguimos o protocolo de análise estatística descritiva e inferencial para variáveis categóricas apresentado por Freitag (2020). Disponível em: <https://rkofreitag.github.io/Categorica.html>. Acesso em: 23 jun. 2020.

**GRÁFICO 2.** Ocorrências do pronome possessivo anteposto versus ocorrências do pronome possessivo posposto ao núcleo nominal em Neves (1993) e Franchi (1996)

$\chi^2_{\text{gor}}(1) = 326.34$ ,  $p = 6.01\text{e-}73$ ,  $\widehat{V}_{\text{Cramer}} = 0.90$ ,  $\text{CI}_{95\%} [0.80, 0.99]$ ,  $n_{\text{obs}} = 406$

$\chi^2_{\text{gor}}(1) = 866.25$ ,  $p = 2.13\text{e-}190$ ,  $\widehat{V}_{\text{Cramer}} = 0.87$ ,  $\text{CI}_{95\%} [0.71, 0.94]$ ,  $n_{\text{obs}} = 1136$



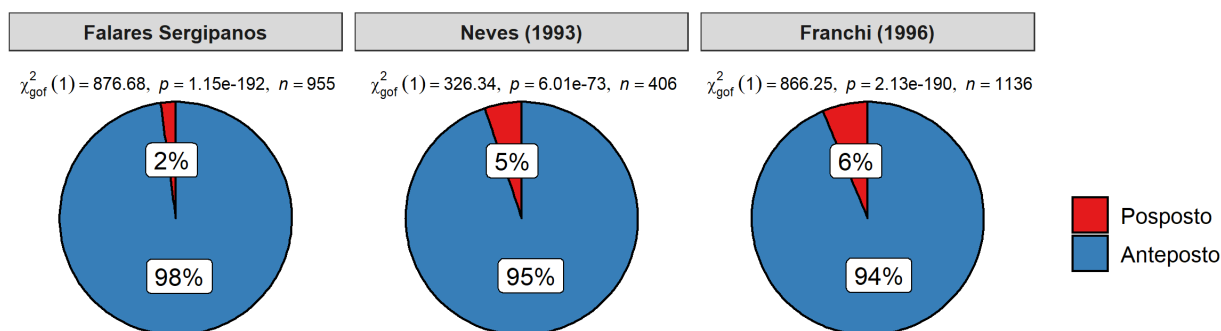
Fonte: Produzido pelo autor, extraído de MÜLLER (1998, p. 12)

Para os dados de Neves (1993), a frequência de possessivos pospostos é menor (5%  $n = 21$ ) do que a frequência de possessivos antepostos (95%  $n = 385$ ), e essa diferença é estatisticamente significativa, confirmada pelo valor de qui-quadrado, em que  $X^2(1, N = 406) = 326.34$ ,  $p < 0,001$ , com associação forte ( $V^2 = 0.90$ ). Em Franchi (1996), a frequência de possessivos pospostos também é menor (6%  $n = 72$ ) do que a frequência de possessivos antepostos (94%  $n = 1064$ ) e essa diferença também é significativa, confirmada pelo valor de qui-quadrado, em que  $X^2(1, N = 1136) = 866.25$ ,  $p < 0,001$ , com associação forte ( $V^2 = 0.87$ ).

As percentagens globais entre os dados de Neves (1993) e Franchi (1996) se aproximam entre si. Para os possessivos antepostos, há 95% ( $n = 385$ ) para Neves (1993) e 94% ( $n = 1064$ ) para Franchi (1996). Na amostra Falares Sergipanos (98%  $n = 935$ ), há diferença de 3% com os dados de Neves (1993). Em todas as amostras, predomina o possessivo anteposto, enquanto os pospostos ocorrem em números bastante baixos. Correlacionamos os três resultados para observar se há associação entre as distribuições das pesquisas (Gráfico 3):

**GRÁFICO 3.** Distribuição das ocorrências nos dados de Neves (1993), Franchi (1996) e Falares Sergipano

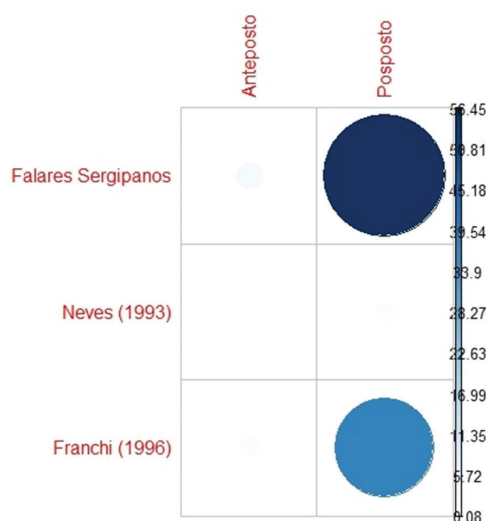
$\chi^2_{\text{Pearson}}(2) = 22.10$ ,  $p = 1.59\text{e-}05$ ,  $\widehat{V}_{\text{Cramer}} = 0.09$ ,  $\text{CI}_{95\%} [0.05, 0.13]$ ,  $n_{\text{obs}} = 2497$



Fonte: Produzido pelo autor

Há diferença estatisticamente significativa entre os dados, confirmada pelo teste de qui-quadrado, em que  $X^2$  ( $N= 2497$ ) = 22.10,  $p < 0,001$ , o que demonstra que a distribuição entre as três pesquisas é independente: a associação entre as amostras é fraca ( $V^2 = 0.09$ ). Com isso, embora os dados estejam associados, o efeito da associação é extremamente baixo, o que nos leva a predizer que não há, de fato, associação entre os dados das pesquisas. Podemos ver o quanto cada amostra contribui para o escore do teste de qui-quadrado por meio de uma matriz de correlação (Gráfico 4).

**GRÁFICO 4.** Contribuição de cada fator para o valor de qui-quadrado quanto a distribuição dos tipos de SN possessivizados



Fonte: Produzido pelo autor

No Gráfico 4, o tamanho do círculo é proporcional à contribuição da célula para o resultado do teste. A posição posposta na amostra Falares Sergipanos é o fator que mais contribui para o resultado, o que mais se diferencia, seguido pela posição posposta na amostra de Franchi (1993). A posição anteposta em Falares Sergipanos, a posposta em Neves (1993) e a anteposta em Franchi (1996) também contribuem, mas a dispersão do círculo demonstra que a contribuição é muito pouca.

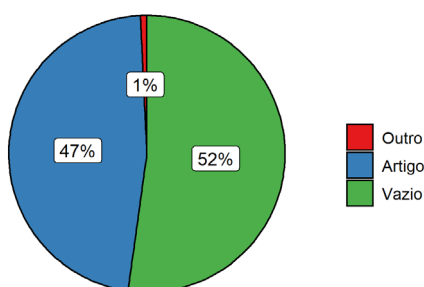
São os números do SN possessivizado com possessivo posposto que apresentam efeito sobre a diferença entre a frequência observada e a frequência esperada, e são os fatores mais distintos nos dados da distribuição dos SN possessivizados em português. Com isso, quanto à distribuição dos SN possessivizados, não há associação entre os dados das pesquisas, o que não permite apresentar uma generalização.

Passemos para os tipos de sintagmas nominais possessivizados, começando pelos possessivos antepostos na amostra Falares Sergipanos (Gráfico 5)<sup>5</sup>, que correspondem a 98% ( $n= 937/955$ ) na amostra.

<sup>5</sup> O leitor pode observar que há diferença entre os elementos apresentados no Gráfico 5 e a organização do SN possessivizado anteposto no Quadro 1: só inserimos três classes de palavras para manter a similaridade com o que é apresentado na pesquisa de Neves (1993) e, assim, poder desenvolver a comparação. O mesmo quanto a organização do SN possessivizado posposto.

**GRÁFICO 5.** Distribuição de elementos à esquerda de possessivo anteposto ao nome no banco Falares Sergipanos

$$\chi^2_{\text{got}}(2) = 451.42, p = 9.43\text{e-}99, \hat{V}_{\text{Cramer}} = 0.49, \text{CI}_{95\%} [0.44, 0.53], n_{\text{obs}} = 937$$



Fonte: Produzido pelo autor

A frequência de possessivos antepostos com a posição vazia, como em (30), é maior (52%  $n=489$ ) do que a frequência dos possessivos antepostos acompanhados de artigos definidos (47%  $n=441$ ), como em (31), e do que acompanhados por outros elementos (1%  $n=7$ ), como em (32) e (33). Essa diferença é estatisticamente significativa, confirmada pelo valor de qui-quadrado, em que  $X^2(2, N=955) = 451.42, p < 0,001$ , com associação média ( $V^2 = 0.49$ ). Nos dados do banco Falares Sergipanos predomina a construção Poss + N, sem nenhum outro elemento no SN possessivizado.

(30) então “meu pai” é motorista e “minha mãe” é servente masculino dezenove Aracaju Sergipe. (BRE1MF)

(31) frente “a minha expectativa” é que “a minha vida” ia ser mais tranquila quando eu (CLA-1FF).

(32) ele ta- “esse meu amigo” ele tava reclamando. (ROD1MI)

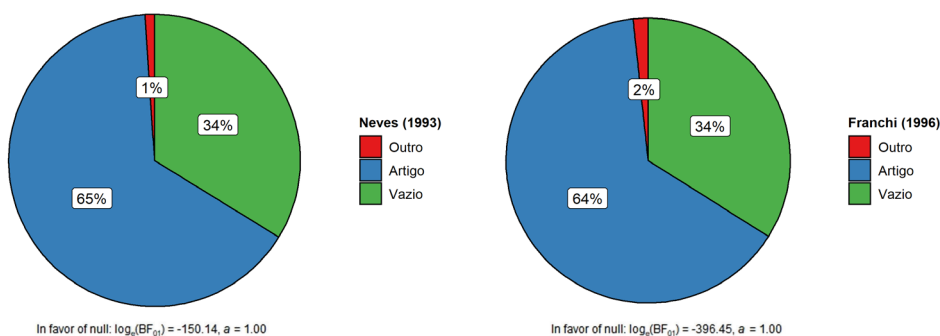
(33) contrataram todos os “todos os meus colegas” no caso. (CAS1MF)

Os resultados encontrados por Neves (1993) e Franchi (1996) são apresentados abaixo (Gráfico 6). Relembremos que os possessivos antepostos em Neves (1993) correspondem a 95% (385/406), e a 94% (1064/1136) em Franchi (1996).

**GRÁFICO 6.** Ocorrência de determinante antes do grupo “pronome adjetivo possessivo” + “nome” em Neves (1993) e em Franchi (1996)

$$\chi^2_{\text{got}}(2) = 237.73, p = 2.39\text{e-}52, \hat{V}_{\text{Cramer}} = 0.55, \text{CI}_{95\%} [0.48, 0.62], n_{\text{obs}} = 385$$

$$\chi^2_{\text{got}}(2) = 627.36, p = 5.89\text{e-}137, \hat{V}_{\text{Cramer}} = 0.54, \text{CI}_{95\%} [0.50, 0.58], n_{\text{obs}} = 1064$$



In favor of null:  $\log_e(\text{BF}_{01}) = -150.14, a = 1.00$

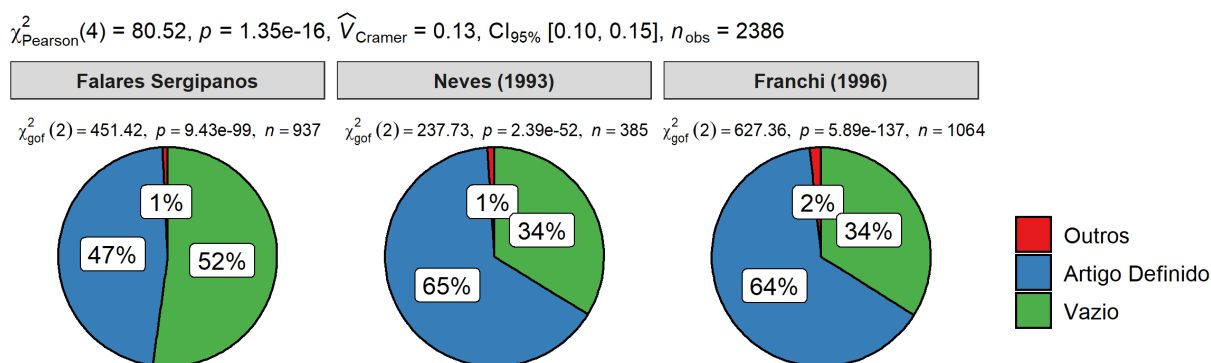
In favor of null:  $\log_e(\text{BF}_{01}) = -396.45, a = 1.00$

Fonte: Produzido pelo autor, extraído de MÜLLER (1998)

A frequência para artigo definido no SN possessivizado é maior nos dados de Neves (1993), com 65% (n= 251), seguindo pela posição vazia, com 34% (n= 130), e outros, com 1% (n= 4). A diferença na frequência é estatisticamente significativa, em que  $X^2(2, N= 385) = 237.73$ ,  $p < 0,001$ , com associação média ( $V^2 = 0.56$ ). Em Franchi (1996), também há maior frequência de artigo definido, com 64% (n= 685), seguido por posição vazia, com 34% (n= 361) e outros elementos, com 2% (n= 18). A diferença na frequência é estatisticamente significativa, em que  $X^2(2, N= 1064) = 627.36$ ,  $p < 0,001$ , com associação média ( $V^2 = 0.54$ ).

Os percentuais apresentados pelos autores apontam que nos dados do NURC o que prevalece é o uso de artigo definido antes de possessivos antepostos. Correlacionamos os resultados das três amostras para observar se há associação (Gráfico 7).

**GRÁFICO 7.** Distribuição das ocorrências de possessivos anteposto nos dados de Neves (1993), Franchi (1996) e Falares Sergipano



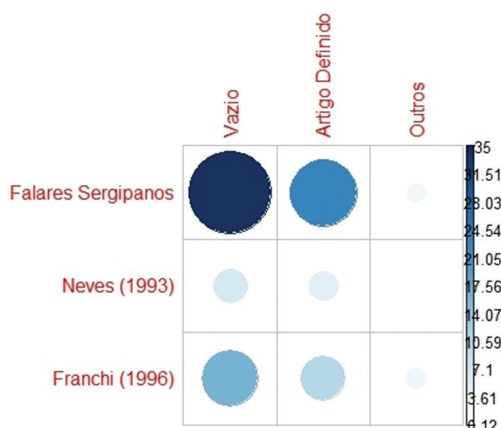
Fonte: Produzido pelo autor

A diferença entre os resultados é estatisticamente significativa, confirmada pelo valor de qui-quadrado, em que  $X^2(N= 2386) = 22.10$ ,  $p < 0,001$ , os dados são independentes. Há associação entre as amostras, em que ( $V^2 = 0.13$ ), o que demonstra que há uma associação fraca entre os dados: mesmo que os dados estejam associados, essa associação é pequena, quase inexistente.

Os testes estatísticos confirmam que nos dados do NURC predomina o uso de artigo definido, enquanto nos dados de Sergipe predomina a posição vazia, o que invalida a generalização de que há uma rigidez na organização dos SN possessivizados antepostos no PB. Essa distinção reflete discussões em relação ao caráter dialetal desses usos, como já sinalizamos em outro trabalho (cf. SIQUEIRA, 2020): os dados do NURC apresentam tendência ao uso do artigo definido na posição, enquanto nos dados de Sergipe e em outras regiões do Nordeste a tendência é para o não preenchimento (cf. SILVA, 1982; 1998a; 1998b; CALLOU e SILVA, 1997; GUEDES, 2019; SEDRINS et al., 2019).

A posição vazia na amostra Falares Sergipanos é o fator que mais contribui para o escore do teste, seguido pela posição com artigo definido na amostra Falares Sergipano: são os fatores que mais diferem dos demais na distribuição e os que mais contribuem para o baixo valor de asso-

**GRÁFICO 8.** Contribuição de cada fator para o valor de qui-quadrado quanto a distribuição dos possessivos antepostos



Fonte: Produzido pelo autor

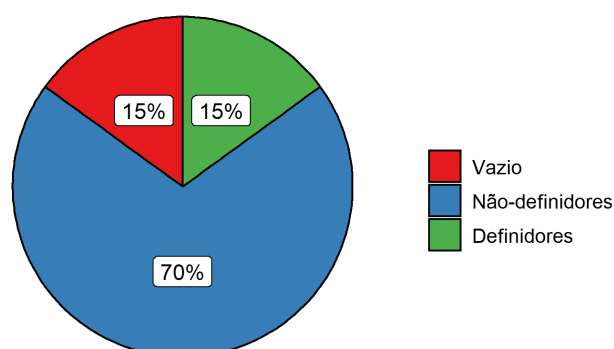
ciação. Os fatores vazio e artigo definido também apresentam contribuição em Franchi (1996) e Neves (1993), porém, a proporção do círculo demonstra que essa contribuição é pouca. A não associação entre os dados quanto à organização do SN possessivizado anteposto é decorrente, principalmente, da distribuição na amostra Falares Sergipano, que difere da distribuição em Neves (1993) e Franchi (1996).

Passemos para os possessivos pospostos, i. e. que ocorrem após o nome, como em (34), que representam 2% (20/955) de nossos dados (Gráfico 9):

(34) nosso Ensino Médio que era “patrimônio nosso” que só tinha aqui (ROD1MI).

**GRÁFICO 9.** Distribuição de elementos à esquerda do possessivo posposto ao nome

$$\chi^2_{\text{gor}}(2) = 12.10, p = 0.002, \hat{V}_{\text{Cramer}} = 0.53, \text{CI}_{95\%} [0.14, 0.79], n_{\text{obs}} = 20$$



Fonte: Produzido pelo autor

Nos dados do banco Falares Sergipanos, o predomínio é de não-definidores (65% n= 13), em (35), seguido por posição vazia (15% n= 3), em (36), e definidores (15% n= 3), em (37), (38) e



(39). A diferença na distribuição das frequências é estatisticamente significativa, em que  $X^2$  (2,  $N=2497$ ) = 12.10,  $p < 0,002$ , com associação média ( $V^2 = 0.55$ ).

(35) que eu perdi “um óculos meu” um óculos de grau. (DAN1FI)

(36) nenhuma aluno quer perder “aula sua” nenhum aluno quer tar longe. (LOR1FI)

(37) ela não tinha “esse tempo meu” também tava trabalhando como motorista. (BRE1MF)

(38) você não pega “o frete seu”. (ROD1MI)

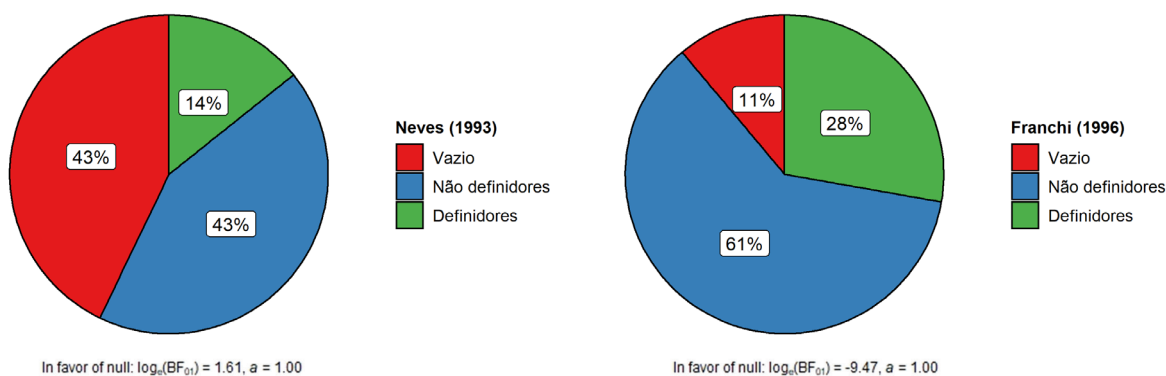
(39) todos eles “oito nove tios meus” só por parte de mãe. (GRE1MF)

Os possessivos pospostos no português sergipano são realizados, preferencialmente, com um artigo indefinido antes do nome. Mesmo que outros elementos ocorram, seu aparecimento é bastante reduzido, levando em consideração que possessivos pospostos já tendem a ocorrer em muito menor número. No Gráfico 10, vemos os resultados obtidos por Neves (1993), nos quais os possessivos pospostos representam 5% (21/406) e Franchi (1996), que representam 94% (72/1136):

**GRÁFICO 10.** Ocorrência de determinante antes do grupo “nome” + “pronome adjetivo possessivo” em Neves (1993) e Franchi (1996)

$\chi^2_{\text{gof}}(2) = 3.43$ ,  $p = 0.180$ ,  $\hat{V}_{\text{Cramer}} = 0.19$ ,  $CI_{95\%} [0.00, 0.44]$ ,  $n_{\text{obs}} = 21$

$\chi^2_{\text{gof}}(2) = 28.00$ ,  $p = 8.32\text{e-}07$ ,  $\hat{V}_{\text{Cramer}} = 0.43$ ,  $CI_{95\%} [0.25, 0.58]$ ,  $n_{\text{obs}} = 72$



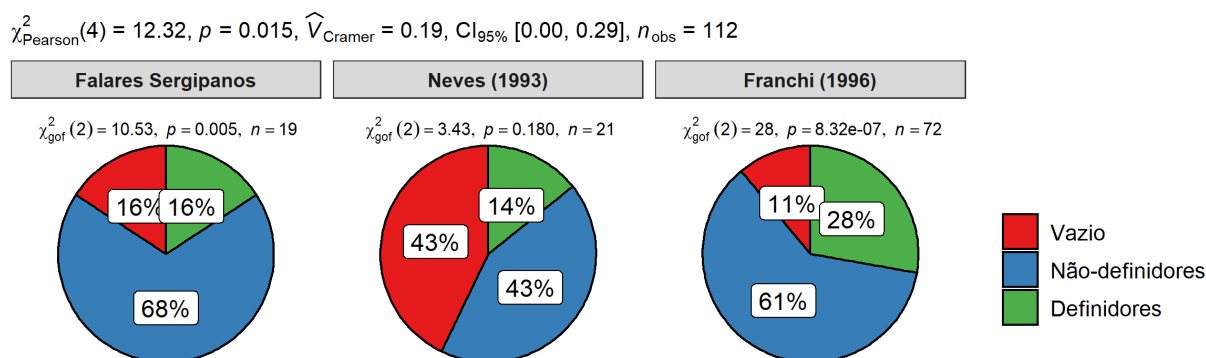
Fonte: Produzido pelo autor, extraído de MÜLLER (1998)

Em Neves (1993), predominam os não-definidores (43%  $n=9$ ) e a posição antes de nome vazia (43%  $n=9$ ). Os definidores ocorrem com baixos números (14%  $n=3$ ). A diferença entre a frequência observada e a frequência esperada não é estatisticamente significativa, confirmada pelo valor que qui-quadrado, em que  $X^2$  (2,  $N=21$ ) = 3.43,  $p = 0.180$ . Em Franchi (1996), também há o predomínio de não-definidores (61%  $n=44$ ), seguido pelos definidores (28%  $n=20$ ) e, por fim, a posição antes de nome vazia (11%  $n=8$ ). Nesses dados, a diferença é estatisticamente significativa, confirmada pelo valor de qui-quadrado, em que  $X^2$  (2, 72) = 28.00,  $p < 0,001$ .

Segundo Müller (1998, p. 19), “a quase total ausência de indefinidos com o pronome anteposto é aqui complementada pela alta porcentagem de ocorrências de indefinidos (mais de 60%)

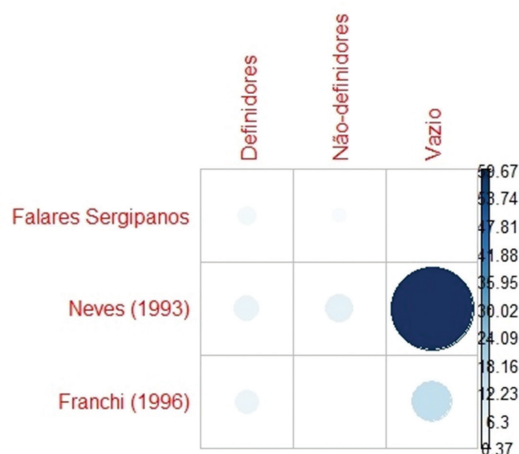
com o pronome possessivo posposto ao núcleo nominal”. Os possessivos pospostos nos dados do NURC são realizados, preferencialmente, com não-definidores: são 44 (61%) realizações de indefinidos nos dados de Franchi (1996) e 9 (43%) em Neves (1993). Para Müller (1998, p. 18), “as percentagens obtidas por ambos são novamente bastante próximas e os dados de Franchi (1996) legitimam as percentagens obtidas por Neves (1993), cujo *corpus* era extremamente restrito no caso dos pronomes possessivos pospostos”. O Gráfico 11 apresenta os dados quanto às três amostras.

**GRÁFICO 11.** Distribuição das ocorrências de possessivos posposto nos dados de Neves (1993), Franchi (1996) e Falares Sergipano



A diferença entre os resultados é estatisticamente significativa, confirmada pelo valor de qui-quadrado, em que  $X^2 (N = 112) = 12.32, p = 0.015$ , os dados são independentes. Há uma associação bem fraca entre as amostras, em que ( $V^2 = 0.19$ ). Mesmo que os dados estejam associados, essa associação é pequena. Os não-definidores predominam em todas as pesquisas, seguidos pelos definidores. A contribuição para o resultado é apresentada a seguir (Gráfico 12).

**GRÁFICO 12.** Contribuição de cada fator para o valor de qui-quadrado quanto a distribuição dos possessivos pospostos



O fator que mais contribui para o resultado do teste é a posição vazia nos dados de Neves (1993), amostra que não apresentou significância estatística. Os definidores apresentam diferença em todas as pesquisas. Mesmo com a proporção dos círculos sugerindo que a contribuição é fraca, demonstra-se que os resultados mais diferem do que parecem entre si. Com isso, diferentemente do que pontuou Müller (1998) quanto aos resultados de Neves (1993) e Franchi (1996) com possessivos pospostos, não há similaridade legitimada entre os dados. A associação que há entre as amostras é bastante fraca: os dados tendem a ser independentes entre si, não havendo, do ponto de vista estatístico, uma associação significativamente forte.

Contudo, devemos considerar que o predomínio de não-definidores com possessivos pospostos é uma particularidade compartilhada. Mesmo que os resultados não estejam associados entre si, podemos considerar que seja um comportamento comum no português, diferentemente da organização dos antepostos.

## 6. Considerações finais

A descrição aqui apresentada mostra que há diferentes tipos de organização em relação à estrutura do SN possessivizado no português sergipano. Primeiro, há forte predominância de possessivos antepostos ao nome, construções do tipo Poss + N, na medida em que construções pospostas, N + Poss, ocorrem em baixos números. Segundo, em possessivos antepostos, as construções mais recorrentes envolvem i) o não preenchimento de elementos antes dos possessivos; e ii) preenchimento por meio de artigo definido, o que remonta a discussões sobre a variação no preenchimento da posição determinante por artigo definido antecedendo possessivos antepostos. Terceiro, em contextos de possessivos pospostos, há predominância de não-definidores antecedendo nomes, enquanto outros elementos ocorrem com números bastante baixos.

Os testes de independência e de associação apontam que há mais diferenças entre as três amostras e seus resultados do que semelhanças em ambos os tipos de possessivos, o que não nos permite desenvolver uma generalização do comportamento desses SN para o português. Nossos dados se distanciam, em partes, dos encontrados por Neves (1993) e Franchi (1996) principalmente quanto à organização dos SN possessivizados antepostos, variando entre o uso ou não de artigo definido, comprovado pela matriz de correlação (Gráfico 4). Nos SN possessivizados pospostos, a maioria dos fatores contribuiu para a não associação. Ainda assim, há uma certa rigidez no comportamento da estrutura posposta, mas só em relação aos não-definidores, já que os outros elementos variam nos resultados.

Mesmo seguindo protocolos de amostragem semelhantes aos do NURC, os resultados da amostra Falares Sergipanos são diferentes, o que nos leva a hipotetizar ser resultado da antiguidade da amostra – os dados do NURC foram coletados ao fim da década de 70, enquanto os dados do Falares Sergipanos foram coletados ao longo da década de 2010 – como também por questões dialetais no quis diz respeito à organização dos antepostos. Destarte, por meio desta descrição, podemos ter uma base da organização dos possessivos no português sergipano, ampliando a descrição a configuração de sua gramática dos possessivos.

## REFERÊNCIAS

- BAILEY, G.; TILLERY, J. Some sources of divergent data in Sociolinguistics. In: FOUGHT, C (Ed.). **Sociolinguistic Variation: Critical Reflections**. New York: Oxford University, 2004, p. 11–30.
- BREZINA, V.; WEILL-TESSIER, P.; MCENERY, A. **#LancsBox** v. 5.x. [software]. 2020. Disponível em: <<http://corpora.lancs.ac.uk/lancsbox/>> Acesso: 23 jun. 2020.
- CALLOU, D.; SILVA, G. M. O. O uso do artigo definido em contextos específicos. In: HORA, D. (Org.). **Diversidade Lingüística no Brasil**. João Pessoa: Idéia, 1997.
- CASTILHO, A. T. **Nova gramática do português brasileiro**. 1. ed. 2ª reimpressão. São Paulo: Contexto, 2012.
- CASTILHO, A.; PRETI, D. **A Linguagem Falada Culta na Cidade de São Paulo: Elocuções Formais**, vol. I. São Paulo: Queirós/FAPESP, 1986.
- FRANCHI, C. **Anotações**, (manuscrito), 1996.
- FREITAG, R. M. K. Banco de dados Falares Sergipanos. **Working Papers em Linguística**, v. 14, n. 1, p. 156-164, 2013.
- FREITAG, R. M. K. Sociolinguística no/do Brasil. **Cadernos de Estudos Lingüísticos**, v. 58, n. 3, p. 445-460, 2016.
- FREITAG, R. M. K. Falares sergipanos. In: **Gelne 40 anos**. São Paulo: Blucher, 2017, p. 119-130. DOI 10.5151/9788580392852-06
- FREITAG, R. M. K. Amostras sociolinguísticas: probabilísticas ou por conveniência?. **Revista de Estudos da Linguagem**, v. 26, n. 2, p. 667-686, 2018.
- FREITAG, R. M. K. Reparos na leitura em voz alta como pistas de consciência sociolinguística. **DELTA**, v. 36, n. 2, p. 1-22, 2020.
- FREITAG, R. M. K. **Variáveis categóricas**. Disponível em: <<https://rkofreitag.github.io/Categorica.html/>>. Acesso em: 28 jun. 2020.
- FREITAG, R. M. K. NURC, um banco de dados sociolinguístico. In: Oliveira Jr., M. (org.). **NURC - 50 anos**. São Paulo: Parábola, 2019, p. 125-134.
- FREITAG, R. M. K.; MARTINS, M. A.; TAVARES, M. A. Bancos de dados sociolinguísticos do português brasileiro e os estudos de terceira onda: potencialidades e limitações. **Alfa**, 56 (3): p. 917-944, 2012.
- FREITAG, R. M. K.; ROST-SNICHELOTTO, C. A. Análises contrastivas: estabilidade, variedade ou metodologia?. **Working Papers em Linguística**, v. 16, n. 1, p. 157-169, 2015.
- GUEDES, S. Emprego do artigo definido em situação de contato dialetal. **Domínios de Linguagem**, v. 13, n. 4, p. 1401-1432, 2019.
- HELLWIG, B.; GEERTS, J. **ELAN: Linguistic Annotator**. Versão 4.4.0. 2013. Disponível em: <[mpi.nl/corpus/manuals/manual-elan.pdf](http://mpi.nl/corpus/manuals/manual-elan.pdf)>. Acesso em 23 jun. 2020.
- ILARI, R (org.). **Gramática do português culto falado no Brasil, vol. 3: palavras de classe aberta**. Editora da UNICAMP, 2014.



ILARI, R. (org.). **Gramática do português culto falado no Brasil, vol. 4:** palavras de classe fechada. Editora da UNICAMP, 2015.

ILARI, R.; NEVES, M. H. M. (orgs.). **Gramática do português culto falado do Brasil, vol. 2:** Classes de palavras e processos em construção. Editora da UNICAMP, 2008.

JUBRAN, C. C. A. S (org.). **Gramática do português culto falado no Brasil, vol. 1:** Construção do texto. Editora da UNICAMP, 2006.

LABOV, W. **Principles of linguistic change, volume 1:** Internal factors. Oxford: Blackwell, 1994.

LABOV, W. **Where does the linguistic variable stop?:** A response to Beatriz Lavandera. Southwest Educational Development Laboratory, 1978

MÜLLER, A. L. O significado da ordem os pronomes possessivos no sintagma nominal. **Revista da ANPOLL**, n. 4, p. 11-37, 1998.

NEVES, M. H. M. Possessivos. In: CASTILHO, A. T. (org.) **Gramática do Português Falado, vol. 3.** Campinas: Editora da UNICAMP/FAPESP, 1993, p. 149-211.

NEVES, M. H. M. Os pronomes. In: ILARI, R.; NEVES, M. H. M. (orgs.). **Gramática do português culto falado do Brasil, vol. 2:** Classes de palavras e processos em construção. Editora da UNICAMP, 2008, p. 507-622.

NEVES, M. H. M. (org.). **Gramática do português culto falado no Brasil, vol. 5:** a construção das orações complexas. Editora da UNICAMP, 2016.

PERINI, M. A. **Gramática descritiva do português.** Ática, 2005.

SEDRINS, A. P. *et al.* A função sintática e o licenciamento de artigos definidos diante de antropônimos e de possessivos pré-nominais. **Domínios de Linguagem**, v. 13, n. 3, p. 1266-1295, 2019.

SILVA, G. M. O. **Estudo da Regularidade na Variação dos Possessivos no Português do Rio de Janeiro.** Tese (Doutorado) - Universidade Federal do Rio de Janeiro, 1984.

SILVA, G. M. O. Realização facultativa do artigo definido diante de possessivo e de patronímico. In: SILVA, G. M. O.; SCHERRE, M. M. P. (Orgs.). **Padrões sociolingüísticos:** análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro. Rio de Janeiro: Tempo Brasileiro, p. 120-145, 1998a.

SILVA, G. M. O. Emprego do artigo diante de possessivo e de patronímico: resultados sociais. In: SILVA, G. M. O.; SCHERRE, M. M. P. (Orgs.). **Padrões sociolingüísticos:** análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro. Rio de Janeiro: Tempo Brasileiro, p. 265-281, 1998b.