

O espraçamento de *você* em contextos sintáticos de complementação e adjunção no alto sertão alagoano

The spreading of *você* in syntactic contexts of complementation and adjunction in the high backwoods of Alagoas, Brazil

Elyne Giselle de Santana Lima Aguiar Vitória¹ 

Pedro Henrique Sousa dos Santos¹ 

¹Universidade Federal de Alagoas, Campus Arapiraca. Arapiraca, AL, Brasil

E-mail: elyne.vitorio@gmail.com

E-mail: pedro.henrique@arapiraca.ufal.br

Editor-chefe

Marcia dos Santos
Machado Vieira

Autor correspondente

Elyne Giselle de Santana Lima
Aguiar Vitória
elyne.vitorio@gmail.com

Recebido: 24/05/2023

Aceito: 08/04/2024

Como citar:

VITÓRIO, Elyne Giselle de Santana Lima Aguiar; SANTOS, Pedro Henrique Sousa dos. O espraçamento de *você* em contextos sintáticos de complementação e adjunção no alto sertão alagoano. *Revista Diadorim*, v.25, n.1, e58942, 2023. doi: <https://doi.org/10.35520/diadorim.2023.v25n1e58942>

Resumo

Nos estudos sociolinguísticos, a segunda pessoa do singular (2PS) apresenta uma vasta descrição na posição de sujeito. Entretanto, na posição de complemento e adjunto, os estudos são escassos, o que mostra a importância de novas investigações acerca do espraçamento das formas do paradigma “*você*” em tal posição, em contraposição ao nulo e aos clíticos. Com isso, esta pesquisa, amparada no aporte teórico-metodológico da Sociolinguística Variacionista, objetivou analisar a variação desses pronomes em uma amostra de 96 informantes coletada pelo projeto LUSA (A Língua Usada no Sertão Alagoano), controlando as variáveis sociais *sexo/gênero, faixa etária e escolaridade*, bem como as variáveis linguísticas *relação gramatical, tipo de núcleo, paralelismo formal e semântico e tipo de referente*. Os dados foram analisados quantitativamente no programa RStudio com o uso de ferramentas para a criação de gráficos, de um modelo de regressão multinomial e de uma árvore de inferências condicionais.

Como resultado, observou-se a baixa presença, na região do alto Sertão, do paradigma “você” e a maior utilização da variante nula, o que não foi observado em nenhuma das pesquisas analisadas, com essa variação sendo condicionada pelas variáveis *relação gramatical, tipo de referente e faixa etária*. Assim, esses resultados contribuem para uma nova perspectiva acerca do fenômeno, mostrando uma possível diferença regional na variação da 2PS em posição de não sujeito na região sertaneja de Alagoas.

Palavras-chave

Sociolinguística; 2PS; Não Sujeito; Sertão Alagoano.

Abstract

In sociolinguistic studies, the second person of singular (2PS) presents a wide description in the subject position. However, in the complement and adjunction positions, studies are scarce, which shows the importance of new investigations about the spreading of the forms of the “você” paradigm in such a position, as opposed to the null and the clitics. Thus, this research, supported by the theoretical-methodological contribution of Variationist Sociolinguistics, aims to investigate the variation of these pronouns in a sample of 96 informants collected by the LUSA project (A Língua Usada no Sertão Alagoano), controlling the social variables *sex /gender, age group* and *education*, as well as the linguistic variables *grammatical relationship, head type, formal and semantic parallelism* and *referent type*. Data were analyzed quantitatively in RStudio using tools for creating graphs, a multinomial regression model and a conditional inference tree. As a result, there was a low presence, in the backwoods region, of the “você” paradigm and a greater use of the null variant, which was not observed in any of the analyzed studies, with this variation being conditioned by the variables *grammatical relationship, referent type* and *age group*. Thus, these results contribute to a new perspective on the phenomenon, showing a possible regional difference of the high backwoods region of Alagoas in the variation of 2PS in non-subject position.

Keywords:

Sociolinguistics; 2PS; Non-subject; Alagoas Backwoods.

Introdução

A variação entre as formas da segunda pessoa do singular (2PS) já vem sendo estudada dentro da Sociolinguística Variacionista há algumas décadas, com trabalhos em diversas variedades do Brasil (MENON, 1995; LOPES, 2007; DUARTE, 2013; SCHERRE, *et al.*, 2015, entre outros). Contudo, nas posições de não sujeito, isto é, de complemento e adjunto, ainda há poucos trabalhos investigando a variação entre as formas dos clíticos, de “você” e do objeto nulo (GAMA, 2018; ALMEIDA, 2014; ARAÚJO; BORGES, 2021, entre outros), como em (1a), (1b) e (1c), formulados por introspecção.

- (1)a. Eu te amo.
- b. Eu gosto de você.
- c. Eu não sei \emptyset responder.

Nesse contexto, o presente trabalho objetivou investigar como o paradigma “você” e suas formas correspondentes (você, prep. + você¹) está implementado no falar da comunidade do alto sertão de Alagoas. Como não há trabalhos sobre a variação da 2PS nas posições de complemento e adjunto nessa região, a pesquisa contribui para a documentação linguística de regiões interioranas, o que, por sua vez, promove o entendimento da estrutura do português brasileiro e das pressões sociais envolvidas na variação e na mudança da língua.

A pesquisa empreendida tomou como aporte teórico-metodológico os pressupostos da Sociolinguística Variacionista (Weinreich; Labov; Herzog, 2006; Labov, 2008), que analisa a língua como sistema inerentemente variável. Essa área desprende-se das introspecções de teorias formalistas da língua, logo, investiga, a partir do uso real da língua, por que ela varia/muda com o passar do tempo. Com isso, essa vertente de estudos propõe correlacionar pressões linguísticas e sociais que atuam no condicionamento de formas linguísticas variantes.

Para analisar a variação da 2PS em posição de não sujeito, recorreu-se ao problema dos fatores condicionantes, com o objetivo de testar se variáveis linguísticas (relação gramatical, tipo de núcleo, paralelismo formal e semântico, e tipo de referente) e sociais (sexo/gênero, faixa etária e escolaridade) condicionam o uso da 2PS nas funções de não sujeito. Para tanto, faz-se uso do programa RStudio (R Core Team, 2022), que usa a linguagem de programação R, a fim de promover diferentes cálculos estatísticos, assim como a feitura de ilustrações.

¹“Prep. + você” se refere à construção “preposição + você”, como em (1b).

O artigo está estruturado da seguinte forma: a seção “Revisão de literatura” descreve os trabalhos sobre o tópico em estudo, mostrando um panorama desse fenômeno no português brasileiro, em seus vieses diacrônico e sincrônico; a seção “Metodologia e hipóteses” descreve a metodologia utilizada na pesquisa e as hipóteses, alternativas e nulas, acerca da 2PS em posição de não sujeito, constituídas a partir da revisão de literatura. Na seção “Resultados e discussões”, apresentam-se os resultados e sua relação com as hipóteses elencadas. Na seção “Considerações finais”, é feito um balanço das contribuições da pesquisa para a área da Sociolinguística Variacionista e para a compreensão da variedade sertaneja alagoana.

Revisão de literatura

Nos estudos normativos, as formas do paradigma “você” na função de sujeito, como em (2), recebem um tratamento marginalizado, segundo Lopes (2007). Nesta seção, investiga-se se o mesmo ocorre em posições de complemento e adjunto, fazendo uso de alguns manuais gramaticais mais consagrados, como Rocha Lima (2011), Bechara (2014a, 2014b, 2019) e Cunha e Cintra (2017), para, em seguida, contrapô-los a descrições dos estudos sociolinguísticos.

(2)**Você** disse que eu te acharia na faculdade para pegar o teu livro. (Lopes, 2007, p. 103, adaptado)

Somente em Cunha e Cintra (2017) foram encontradas menções à forma “você” em posição de complemento e adjunto. Nessa gramática, os autores explicitam que “você” pode ser usado em funções de objeto direto e indireto, agente da passiva e adjunto, substituindo as usuais formas átonas. A exclusão dos outros manuais pode se dar, no caso do “tu”, por ser uma forma mais conservadora e mais marcada, o que restringe o uso desse pronome em posição de não sujeito; e, no caso do “você”, por não ser reconhecido pelas gramáticas normativas no quadro dos pronomes pessoais do caso reto, logo, não tendo sua contraparte como argumento interno do verbo ou adjunto.

Já nas gramáticas descritivas (Castilho, 2010; Neves, 2011; Perini, 2016), uma nova perspectiva de análise se abre sobre o fenômeno aqui investigado. Perini (2016) aponta que “tu” varia com “você” na posição de sujeito e possui a forma oblíqua “te” como contraparte, mas que também pode ser substituída pelo pronome reto tanto em posição de objeto direto (*eu amo você/ eu te amo*) quanto indireto (*Eu vou te contar/ Eu vou contar para você*). Já em Neves (2011), há uma descrição apenas da forma “você” reta ocorrendo em posições “não sujeito”, sem colocá-la no quadro dos pronomes pessoais oblíquos. Em Castilho (2010), reconhece-se o uso de “você” em

posições de complemento no português brasileiro informal, não estando presente na variedade formal da língua.

Dessa forma, pode-se observar que as gramáticas descritivas capturam o fenômeno da forma “você” em posição de não sujeito com mais clareza do que as gramáticas normativas. Entretanto, nem uma nem outra mencionam o uso do pronome “tu” nessa posição, o que pode indicar a restrição dessa forma linguística a uma certa região ou ocorrência em comunidades ainda não documentadas pelos estudos linguísticos. Logo, a análise de materiais descritivos e normativos ajudam a traçar hipóteses do comportamento desse fenômeno linguístico variável no *corpus* a ser analisado na presente pesquisa.

Nos estudos sociolinguísticos, há uma ampla literatura que analisa a variação “tu” e “você” em posição de sujeito (Menon, 1995; Lopes, 2007; Lopes, 2012; Duarte, 2013; Scherre, *et al.*, 2015; Galves *et al.*, 2016; Yacovenco; Scardua, 2017; Silva; Vitória, 2017), tanto na sincronia como na diacronia do português brasileiro. Os autores reconhecem, em convergência, que a inserção de “você” na posição de sujeito desencadeou mudanças na organização do sistema pronominal do português, o qual ficou dividido em três subsistemas concorrentes: i) “tu” exclusivo; ii) “você” exclusivo”; e iii) “tu” e “você” (Galves *et al.*, 2016). Nos subsistemas 2 e 3, em posições acusativas, dativas e oblíquas (Duarte, 2003), as formas advindas do paradigma “você” ganharam espaço, apesar de, como será visto, a posição acusativa ainda continuar sendo mais restritiva às formas inovadoras.

Já em posições de não sujeito, as formas inovadoras não se difundiram com a mesma regularidade (Lopes *et al.*, 2018). Nas pesquisas analisadas, a forma “tu” não sujeito é brevemente citada (Rumeu, 2015; Araújo; Borges, 2021) em favorecimento de sua forma “te” muito produtiva, como em (1a), embora os estudos diacrônicos foquem mais em quais fatores influenciam o uso das formas do paradigma “você” em complemento e adjunto. As duas subseções seguintes reportam como as pesquisas diacrônicas e sincrônicas analisaram a difusão dessas expressões.

Os estudos diacrônicos

A análise da distribuição regional do uso do paradigma “você” no português brasileiro encontra resultados similares nos trabalhos analisados. Cruz e Rumeu (2016) reportam que, na diacronia mineira, há uma prevalência das formas do paradigma “você” (71,8%, n = 145) sobre as do “tu” (18,3%, n = 37) e do nulo (9,9%, n = 20), o que é corroborado por Rumeu (2015), que, por outro lado, também apresenta resultados para cartas do Rio de Janeiro, em que houve mais formas do paradigma “tu” (59%, n = 160), divergindo de Minas Gerais. Rumeu e Oliveira (2016) também apontam para a prevalência das formas de “tu” sobre as de “você” na variedade do

Rio de Janeiro no século XX, com 76% das ocorrências (n = 345) para a primeira, 19% (n = 87) para a última e 5% (n = 24) para o nulo.

Na análise de vários estados feita por Lopes *et al.* (2018), atesta-se que estados nordestinos, como Pernambuco e Bahia, apresentam maior uso das variantes inovadoras, enquanto, na região Sul e em alguns estados do Sudeste, a forma conservadora permanece. Entretanto, vale ressaltar que Lopes *et al.* (2018) apresentam os dados subdivididos por relação gramatical, que será analisada a seguir, o que leva a uma dificuldade de interpretação geral da distribuição das formas de “tu” e “você” no português brasileiro dos séculos XIX e XX.

No estudo da variável *relação gramatical*, as pesquisas diacrônicas convergem nos resultados. A relação acusativa, objeto direto do verbo, apresenta mais conservadorismo na diacronia analisada em Rumeu (2015) e Rumeu e Oliveira (2016), com a prevalência do “te” no Rio, em cartas mineiras do século XX (Rumeu, 2014; Cruz; Rumeu, 2016) e no Nordeste, no Sudeste e no Sul do Brasil (Lopes *et al.*, 2018). Já, nas relações dativas, objeto indireto do verbo que pode se cliticizar em “lhe”, e oblíquas, posições regidas por preposição que não podem ser cliticizadas em “lhe”, as formas de “você” ganham mais espaço para se encaixar no sistema linguístico, com destaque para o oblíquo, no qual prevalece a forma “preposição + você”. Além disso, as pesquisas vêm mostrando que essa relação gramatical é a que mais favorece o objeto nulo (Rumeu, 2015; Cruz; Rumeu, 2016; Lopes *et al.*, 2018). Nesse sentido, espera-se que novos *corpora* analisados vão ao encontro dos estudos abordados, já que há concordância entre eles quanto à influência dessa variável.

Quanto à variável *tipo de núcleo*, que analisa o núcleo ao qual se liga o sintagma das relações gramaticais já abordadas acima, os resultados também convergem para uma mesma conclusão. Em Rumeu (2014), evidencia-se que o núcleo verbal é um contexto suscetível à entrada do “você” no sistema pronominal do português brasileiro na relação dativa (52%, n = 67), enquanto, na relação oblíqua, o “você” foi mais produtivo ligado a um núcleo nominal (83%, n = 60) — o que pode ser observado, respectivamente, nos exemplos (3a) e (3b).

- (3)a. “[...] eu escrevi para **você** mas não pude mandar [...]” (Rumeu, 2014, p. 96, adaptado)
b. “[...] estou muito grato **a você** pela dádiva.” (Rumeu, 2014, p. 98, adaptado)

Já em Rumeu e Oliveira (2016), as formas inovadoras do paradigma “você” apareceram com maior frequência ligadas a núcleo verbal (74%, n = 65), assim como o objeto nulo — todos os dados estavam ligados a núcleo verbal (n = 24). Do mesmo modo, Cruz e Rumeu (2016) concluíram que os núcleos verbais são os gatilhos

principais para o acionamento da forma inovadora em posição dativa (83%, n = 136), enquanto “você” não sujeito na relação oblíqua, seja em complemento (79%, n = 30) ou adjunto (21%, n = 8), foi realizado, na maioria dos dados, ligado a um nome. É possível assumir que o nível oracional é a porta de entrada mais fácil para o “você” não sujeito, enquanto, no oblíquo, o “você” adentra pelo nível nominal.

Outra variável analisada nos trabalhos foi o *Paralelismo Formal e Semântico*, que investiga se a precedência da forma “você” numa sequência discursiva motiva a presença do “você” em complementos e adjuntos. Rumeu (2014) demonstra que o paralelismo “você_{sujeito}” - “você_{não sujeito}” acontece em 46% dos dados encontrados (n = 92), com em (4a), mas “você_{não sujeito}” - “você_{não sujeito}” é menos frequente (16%, n = 33), como em (4b), revelando que elementos inovadores como sujeito são contextos difusores de formas inovadoras em posições de não sujeito.

- (4)a. “[...] Arlindo **você** mesmo não imagina como eu fiquei impercionada com aquela notícias do seu amigo N... ter ficado mal **com você** por minha causa [...]” (Rumeu, 2014, p. 107, adaptado)
- b. “Lembrei-me **de você**, que se referia a ele com simpatia [...] Enquanto estou escrevendo **a você**, meu neto faz estripulias pela casa. [...]” (Rumeu, 2014, p. 107, adaptado)

Essas conclusões divergem das de Rumeu e Oliveira (2016), que não encontraram esse paralelismo das formas de “você” na amostra do Rio, indicando uma diferença regional referente ao uso dos pronomes em complementos e adjuntos. Também em Cruz e Rumeu (2016) não foi encontrado paralelismo na análise das missivas mineiras do século XX, em que apenas 11% dos dados são referentes ao paralelismo com “você_{sujeito}” e 5,5% com “você_{não sujeito}”. Assim, verifica-se que não há uma convergência nos resultados analisados, dificultando o vislumbre da real efetividade do paralelismo no favorecimento de determinado pronome em posições de não sujeito.

Na análise dos aspectos sociais, as pesquisas diacrônicas são de uso mais escasso, visto que as variáveis controladas são específicas da investigação da mudança linguística. Nesse sentido, só Rumeu e Oliveira (2016) será utilizada para a análise do fator *sexo/gênero*. Nessa pesquisa, a análise dos textos escritos de um casal carioca revelou que o noivo (homem) produziu menos as formas do paradigma “você” (7%, n = 17), enquanto o uso delas foi favorecido na linguagem da noiva (32%, n = 70), apesar de ela ainda produzir mais formas conservadoras (65%, n = 143). Esses resultados levam os autores a conjecturarem que, no século XX, as mulheres, por terem, em geral, menos acesso à escolaridade do que os homens, produziam formas menos conservadoras (mais distantes do padrão normativo exigido). Desse modo, faz sentido encontrar maior uso de termos do paradigma “você” nas cartas da noiva.

Os estudos sincrônicos

Na fala sergipana, Araújo e Borges (2021) reportaram que 42% dos dados (n = 173) são da forma “preposição + você”, enquanto 21% (n = 87) são realizações de “te”, o que indica a predominância, nessa região, das formas inovadoras. Nessa comunidade, o nulo apresenta baixo uso (13%, n = 55). Já na Bahia do século XXI, destaca-se uma variação entre o uso de “lhe” e “te”, com leve prevalência do primeiro, nos pronomes-objeto, numa análise com amostras de Salvador (Almeida, 2014, 2016), Feira de Santana (Gama, 2018) e de Santo Antônio de Jesus (Almeida, 2014). Em relação a “você”, as pesquisas mostram um desfavorecimento desse paradigma, não chegando a 10% de uso em Salvador e Santo Antônio de Jesus.

Os percentuais do nulo de diferentes análises mostram que é uma variante pouco favorecida nas variedades nordestinas, com 17,7% (n = 158) e 21% em Almeida (2014) para as localidades de Salvador e Santo Antônio de Jesus, o que é corroborado por Almeida (2016), que apresenta pouca distribuição da variante em Salvador, com 12%, e Santo Antônio de Jesus, com 15,5%. No Sul, Dalto (2002) analisa amostras do VARSUL de Florianópolis, Curitiba e Porto Alegre e destaca o uso da forma “te”, sem presença do “tu” não sujeito. Ademais, o objeto nulo parece ser pouco usado no Sul do Brasil nessa amostra, com 76 realizações em 517 dados (Dalto, 2002).

Na variável *Relação Gramatical* analisada por esses trabalhos, a relação acusativa apresenta mais conservadorismo na fala atual do Sul (Dalto, 2002) e na fala universitária sergipana (Araújo; Borges, 2021). Por outro lado, a Bahia apresenta menos conservadorismo, com uma presença forte de “lhe” na função acusativa variando com “te” (Almeida, 2014, 2016; Gama, 2018), embora o uso de “você” seja desfavorecido nas comunidades analisadas. Já nas relações dativas, as formas de “você” apresentam apenas 4,5% dos dados em Santo Antônio de Jesus (Almeida, 2014) e são favorecidas em Sergipe, com 89% (n = 170), conforme Araújo e Borges (2021). Além disso, as pesquisas vêm mostrando que essa relação gramatical é a que mais favorece o objeto nulo (Almeida, 2014).

Na variável *Paralelismo Formal e Semântico*, Araújo e Borges (2021) reportam que o pronome lexical é favorecido pelo paralelismo (58%, n = 22), enquanto o clítico é favorecido quando aparece isolado (41%, n = 143); o objeto nulo, na análise dos autores, teve maior uso (40%, n = 12) quando não havia paralelismo. Almeida (2014), na análise da fala baiana, também reporta que houve paralelismo “*lhe*_{não sujeito}”-“*lhe*_{não sujeito}” (90,5% dos dados em comparação a “te”) e “*você*_{sujeito}”-“*lhe*_{não sujeito}” (85% dos dados em comparação a “te”), indicando uma preferência pelas formas inovadoras.

Na análise dos aspectos sociais, Araújo e Borges (2021), Almeida (2014, 2016) e Gama (2018) oferecem resultados para as variáveis *sexo/gênero*, *faixa etária* e *escolaridade* (tratado como “período do curso” na análise sergipana). Essas pesquisas são melhores representadoras da realidade do português brasileiro, visto que as pressões

sociais tendem a se modificar ao longo do tempo junto ao avanço da sociedade, o que torna necessário tomar como base para as hipóteses pesquisas sobre o momento presente.

Em Araújo e Borges (2021), estudantes universitários sergipanos homens produzem mais a forma inovadora do que a conservadora (33% de uso do clítico pronominal, $n = 75$, contra 56% de uso do pronome lexical, $n = 125$), enquanto o oposto acontece com as mulheres (45% de uso do clítico pronominal, $n = 85$, contra 39% de uso do pronome lexical, $n = 75$). Além disso, em relação ao uso do objeto nulo, foi reportado que as mulheres o utilizaram mais vezes (16% dos dados) do que os homens (11% dos dados). Entretanto, visto que o valor do V^2 de Cramer é baixo (0,15), uma análise de mais dados é necessária para saber se essa distribuição é aleatória ou não. Em Almeida (2014, 2016), também se reporta que os homens usam a forma inovadora “lhe” com maior frequência, enquanto as mulheres preferem a forma conservadora “te”, embora a autora reconheça que, na variedade baiana, a forma “lhe” apresente um traço [+formal/cortês], como será analisado a seguir.

Na variável *faixa etária*, as pesquisas baianas são unânimes ao destacar a preferência da faixa etária 3 pela forma “lhe” e da faixa etária 1 pela forma “te” (Almeida, 2014, 2016; Gama, 2018), embora, entre os jovens, haja um equilíbrio relativo no uso das variantes. Pode-se estranhar o fato de a variante inovadora ser a forma mais usada por idosos, mas Almeida (2016) avalia que isso se dá por essa variante apresentar um traço estilístico/pragmático [+formal], indicando uma relação de poder sobre os interlocutores, enquanto “te” é [+informal], indicando maior solidariedade. Logo, vê-se que, embora seja gramaticalmente inovadora, uma vez que o “lhe” deve ser normativamente utilizado para se referir à terceira pessoa, essa forma é estilisticamente mais formal na variedade baiana.

Na análise da variável *escolaridade*, Araújo e Borges (2021) reportaram que o período do curso analisado (início, meio ou fim) não interfere no uso da variante inovadora ou conservadora: nos três casos analisados, os estudantes produziram mais vezes o pronome lexical em posição de complemento e adjunto do que o clítico pronominal, indicando uma preferência de universitários sergipanos por formas inovadoras. Entretanto, a análise do objeto nulo mostrou que, à medida que avança no curso, o estudante tende a diminuir seu uso do objeto nulo, com 14% ($n = 26$) no início, 15% ($n = 16$) no meio e 10% ($n = 13$) no fim. Em Gama (2018), por outro lado, falantes de Feira de Santana (BA) com menos escolaridade preferem a forma inovadora “lhe”, enquanto, em falantes universitários, há um equilíbrio de uso entre “lhe” e “te” na relação acusativa, que foi a investigada no trabalho, o que é constatado também por Almeida (2014). A autora, além disso, compara as realizações de clíticos com a do pronome tônico “você” e reporta que falantes com ensino

superior usam mais o clítico e falantes sem ensino superior, o pronome “você” em posições de não sujeito.

Assim, estudos sociolinguísticos contribuem para a formulação de hipóteses mais coerentes sobre o fenômeno em estudo, além de estabelecerem variáveis linguísticas e sociais influenciadoras da variação dos clíticos, das formas de “você” e do objeto nulo em posições de complemento e adjunto. No mesmo rumo, o presente trabalho visa a investigar essa variação de forma sincrônica pela análise de uma amostra de fala de 96 informantes da região do alto sertão alagoano, estratificada em *sexo/gênero*, *faixa etária* e *escolaridade*, a fim de perceber como se comporta tal fenômeno linguístico variável nessa variedade.

Metodologia e hipóteses

A presente pesquisa recorre ao aporte da Sociolinguística Variacionista (Weinreich; Labov; Herzog, 2006; Labov, 2008), que teoriza sobre estudo da variação e mudança correlacionando variáveis linguísticas e sociais, postulando que a variação está presente no próprio sistema linguístico, não fora dele, de modo que a variação não é um caos, mas um processo ordenado e, portanto, passível de estudo científico padronizado. Nesse sentido, essa teoria visa a entender os mecanismos condicionantes desses processos naturais da língua por meio de metodologia quantitativa de pesquisa (Mollica; Braga, 2020), fazendo uso de programas estatísticos e *corpora* de dados linguísticos tanto falado quanto escrito.

A amostra utilizada nessa pesquisa faz parte do grupo de pesquisa LUSA (A Língua Usada no Sertão Alagoano) e tem como objetivo mapear o modo como falam os habitantes nascidos e criados na região do Sertão de Alagoas, constituído por 26 municípios e povoado por quase 170 mil pessoas (Vitório, 2020). Essa região é dividida em Médio e Alto Sertão, este último sendo o foco da amostra. Segundo dados do IBGE, o Alto Sertão tem economia majoritariamente agrícola e sua população urbana é de 46,05%, o que configura o espaço como um ambiente diferente dos grandes centros urbanos, como a capital Maceió.

Para construir o perfil social da comunidade de fala, estabeleceu-se como variáveis sociais *sexo/gênero*, *faixa etária* e *escolaridade*. Para preencher tais células, foi construída uma amostra com 96 entrevistas sociolinguísticas (24 células x 4 informantes por célula) do tipo DID. A entrevista foi constituída por uma narrativa, por parte do falante, de experiências pessoais pelas quais passou, pois, conforme Labov (2008), o informante se envolve nesse tipo de narrativa como se estivesse revivendo o momento. Desse modo, é possível “olhar para língua em uso nessa região pelo viés da Sociolinguística Variacionista” (Vitório, 2020, p. 171).

Para a análise do fenômeno, estabeleceu-se como Variável Dependente (VD) *Pronome*, constituída pelas variantes “clítico” (te, lhe), “você” (“você, prep. + você”) e “objeto nulo”. Já como Variáveis Independentes (VI), foram selecionadas variáveis linguísticas e variáveis sociais. As VIs linguísticas estão divididas em *Relação Gramatical* (Acusativo, Dativo e Oblíquo), *Tipo de Núcleo* (Não Verbal e Verbal), *Paralelismo Formal e Semântico* (Há Paralelismo, Não Há Paralelismo, Ocorrência Isolada e Primeiro da Série) e *Tipo de Referente* (Determinado e Indeterminado). Já as VIs sociais são *Sexo/gênero* (Masculino e Feminino), *Escolaridade* (E0 (sem escolarização), E1 (ensino fundamental), E2 (ensino médio) e E3 (ensino superior)) e *Faixa Etária* (F1 (18 a 29 anos), F2 (30 a 44 anos) e F3 (acima de 44 anos)).

Em relação às hipóteses alternativas do trabalho, espera-se que: i) as formas inovadoras, de “você”, tenham maior produtividade no *corpus* analisado, visto que são formas produtivas nos trabalhos que investigaram a fala de comunidades nordestinas; ii) a produtividade das formas inovadoras e do objeto nulo será maior em estruturas dativas e oblíquas, visto que a revisão de literatura apontou que os dativos e, principalmente, os oblíquos são menos restritivos à entrada de formas inovadoras no português brasileiro; e iii) as formas inovadoras e o objeto nulo aparecerão com mais produtividade ligadas a núcleos verbais na relação dativa e a não verbais na relação oblíqua, conforme os dados de Rumeu (2014) e Cruz e Rumeu (2016).

Além dessas, também se espera que: iv) haverá paralelismo entre formas inovadoras e formas conservadoras, assim como entre formas conservadoras e conservadoras, como mostrado em alguns dos textos da revisão de literatura. Por fim, vale ressaltar que não há, nos trabalhos reportados sobre a variação dos pronomes-objeto, o uso da variável *Tipo de Referente*. Entretanto, em pesquisas na posição de sujeito, “você” é mais usado em situações nas quais o referente é indeterminado — aparecendo até em construções existenciais, nas quais a referência é genérica (Lopes, 2007). Logo, espera-se a mesma indeterminação para a posição de não sujeito.

Para as variáveis sociais, espera-se que: i) conforme Araújo e Borges (2021) e Almeida (2016), haverá um percentual maior de formas inovadoras produzido por homens do que por mulheres; ii) em relação à escolaridade, o uso da forma inovadora seja estável entre os falantes com ensino superior e mais usado por falantes com menos escolaridade (Gama, 2018; Araújo; Borges, 2021); e iii) o uso da forma inovadora seja preferível pela faixa etária 3, devido a seu caráter pragmaticamente formal, e equilibrado na faixa etária 1, com percentuais próximos entre as formas conservadoras e inovadoras (Almeida, 2014, 2016; Gama, 2018).

Já em relação às hipóteses nulas, espera-se que: i) haja estabilidade entre as formas conservadoras e inovadoras; e ii) não haja relação entre as variáveis independentes no condicionamento dos pronomes de 2PS em posição de não sujeito na amostra alagoana.

Para confirmar ou refutar as hipóteses elencadas e, dessa forma, atingir os objetivos da pesquisa, a análise estatística dos dados foi feita na plataforma RStudio (R Core Team, 2022), que usa a linguagem de programação R, e promoveu testes de significância, como qui-quadrado, feitura de gráficos, com o pacote *ggstatsplot* (Patil, 2018), e ajuste de um modelo de regressão multinomial (Ripley; Venables, 2002), além de uma árvore de inferência condicional (Zeileis; Hothorn, 2015). Tal procedimento objetivou formular um panorama quantitativo de como está se implementando o paradigma “você” na variedade analisada.

$$\chi^2_{\text{gof}}(2) = 336.03, p = 1.08e-73, \hat{C}_{\text{Pearson}} = 0.74, \text{CI}_{95\%} [0.70, 1.00], n_{\text{obs}} = 284$$

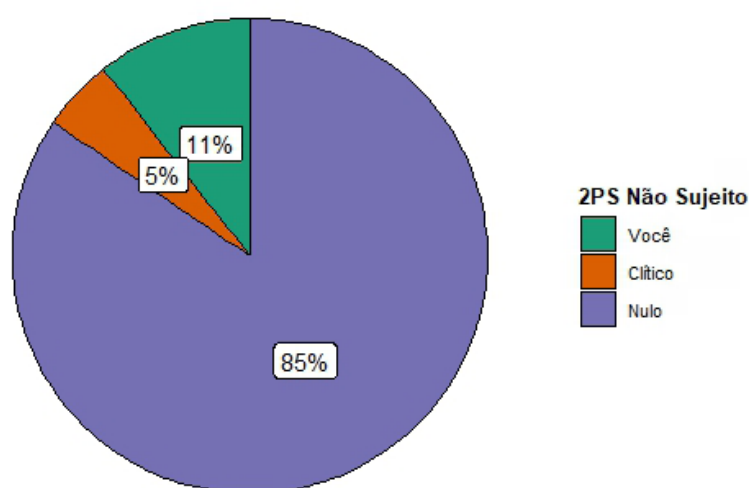


Gráfico 1 – Distribuição das formas pronominais de segunda pessoa em posição de não sujeito

Fonte: elaborado pelos autores.

Resultados e discussões

Na análise dos dados, foram encontradas 284 realizações dos pronomes de 2PS em posição de não sujeito, exemplificadas em (5), (6), (7) e (8). Destas, 240 foram da forma nula (85%), 14, das formas de “clítico” (5%) e 30, das formas de “você” (11%), conforme o Gráfico 1. Esses resultados mostram-se significativos [$\chi^2(2) = 336,03$ $p < 0,001$] e revelam a preferência pela variante nula como referência à segunda pessoa do singular na comunidade, não corroborando a hipótese inicial acerca da preferência da comunidade pelas formas de “você”.

- (5) “aí eu num sei dizê Ø nada” (L83).
(6) “o bom dii você trabalhá **pra você** é isso” (L71).
(7) “Não sei nem **te** dizê x com relação à violência” (L31).²
(8) “então eu não **lhe** aconselho a seguir essa mesma carreira” (L20).

Observando que a preferência pelo nulo não ocorre em outros estados nordestinos, como as pesquisas revisadas anteriormente, é possível que esse comportamento linguístico seja idiossincrático à variedade sertaneja de Alagoas, embora mais trabalhos precisem verificar tal hipótese. Também é necessário investigar a influência do tipo de entrevista sociolinguística nesses resultados, tendo em vista que estudos têm mostrado que esse tipo de coleta não favorece as realizações que apresentam mais proximidade e intimidade entre os falantes.

Como há apenas uma conversa entre o informante e o entrevistador, é possível que haja um certo traço estilístico/pragmático de [+formalidade], visto que há um distanciamento entre os dois membros da entrevista, de modo a favorecer variantes mais respeitadas. Essa hipótese é interessante, pois encontra base na pesquisa de Almeida (2016), que mostra o traço [+informal] da variante “te” e [+formal] de “lhe”. Talvez a variante nula também possa ser [+formal], pois é a variante com mais dados, enquanto “te” só apareceu uma única vez.

Em relação às variáveis independentes, foram feitos testes de qui-quadrado em tabelas nas quais constavam, nas linhas, a VD e, nas colunas, cada VI. Assim, mostraram-se significativas apenas *Tipo de Referente* ($p < 0,001$), *Faixa Etária* ($p < 0,05$) e *Relação Gramatical* ($p < 0,001$) no condicionamento da 2PS em posição de não sujeito. Essas variáveis serão exploradas com mais detalhamento à frente. As variáveis *Sexo/gênero*, *Escolaridade*, *Tipo de Núcleo* e *Paralelismo Formal e Semântico*, conforme Tabela 1, não foram significativas.

A não significância das variáveis sociais *Sexo/gênero* e *Escolaridade* traça uma realidade interessante sobre o uso desses pronomes em Alagoas. Contrariando as hipóteses iniciais com base nos trabalhos já realizados no Nordeste, o não condicionamento traz à luz uma hipótese de que esses pronomes já se espalharam pela comunidade em um processo em que o nulo parece estar vencendo a competição de formas de representação do pronome de 2PS em posição de não sujeito. Assim, não se diferenciam mais os usos de homens e mulheres, independentemente da escolaridade, indicando uma possível não estigmatização das formas linguísticas, de modo que possam ser consideradas variantes que ocorrem nas diversas camadas sociais.

² Nesse único dado, percebe-se que o “te” não foi usado de acordo com a norma-padrão, pois deveria estar em posição acusativa, e, no dado coletado, está na relação dativa (alguém diz algo *a alguém*). Outros aspectos dessa ocorrência são que foi dito por uma pessoa feminina da faixa etária 1 e com nível superior, além de ser uma ocorrência isolada ligada a núcleo verbal e tendo um referente determinado na sequência discursiva.

Tabela 1 – Distribuição dos dados dos pronomes de 2PS em posição de não sujeito em relação às variáveis não significativas

Variáveis	Fatores	Clítico Pronominal	Nulo	Você	Total
SEXO/GÊNERO [χ^2 (2) = 2,5592 $p < 0,2782$]	Masculino	8 (57,1%)	105 (43,7%)	17 (56,6%)	130
	Feminino	6 (42,9%)	135 (56,2%)	13 (43,3%)	154
	Total	14	240	30	284
ESCOLARIDADE [χ^2 (6) = 2,1043 $p < 0,9099$]	E0	3 (21,4%)	51 (21,2%)	4 (13,3%)	58
	E1	2 (14,3%)	42 (17,5%)	4 (13,3%)	48
	E2	4 (28,6%)	79 (32,9%)	12 (40%)	95
	E3	5 (35,7%)	68 (28,3%)	10 (33,3%)	83
Total	14	240	30	284	
TIPO DE NÚCLEO [χ^2 (2) = 1,4096 $p < 0,4942$]	Verbal	14 (100%)	232 (96,6%)	28 (93,3%)	274
	Não Verbal	0 (0%)	8 (3,3%)	2 (6,6%)	10
	Total	14	240	30	284
PARALELISMO FORMAL E SEMÂNTICO [χ^2 (6) = 12,084 $p < 0,06012$]	Ocorrência Isolada	3 (21,4%)	86 (35,8%)	4 (13,3%)	93
	Primeiro da Série	2 (14,2%)	48 (20%)	6 (20%)	56
	Sem Paralelismo	1 (7,1%)	31(12,9%)	8 (26,6%)	40
	Com Paralelismo	8 (57,1%)	75 (31,2%)	12 (40%)	95
Total	14	240	30	284	

Fonte: elaborada pelos autores.

Na relação entre o uso dos pronomes e o tipo de referente utilizado, o Gráfico 2 mostra dados interessantes. Com resultados estatisticamente significativos [χ^2 (2) = 32,64 $p < 0,001$] e a relação entre as variáveis se mostrando moderada, como aponta o valor do V^2 de Cramer (0,33), percebe-se a preferência pela variante nula (n = 211) para se referir a um ente determinado no discurso, enquanto as formas do paradigma “você” e os clíticos são mais favorecidos com referência indeterminada, com o primeiro passando de 7% dos dados (n = 16) para 28% (n = 14) e o segundo, de 3% (n = 7) para 14% (n = 7).

Na relação entre o uso dos pronomes e o tipo de referente utilizado, o Gráfico 2 mostra dados interessantes. Com resultados estatisticamente significativos [$\chi^2(2) = 32,64$ $p < 0,001$] e a relação entre as variáveis se mostrando moderada, como aponta o valor do V^2 de Cramer (0,33), percebe-se a preferência pela variante nula ($n = 211$) para se referir a um ente determinado no discurso, enquanto as formas do paradigma “você” e os clíticos são mais favorecidos com referência indeterminada, com o primeiro passando de 7% dos dados ($n = 16$) para 28% ($n = 14$) e o segundo, de 3% ($n = 7$) para 14% ($n = 7$).

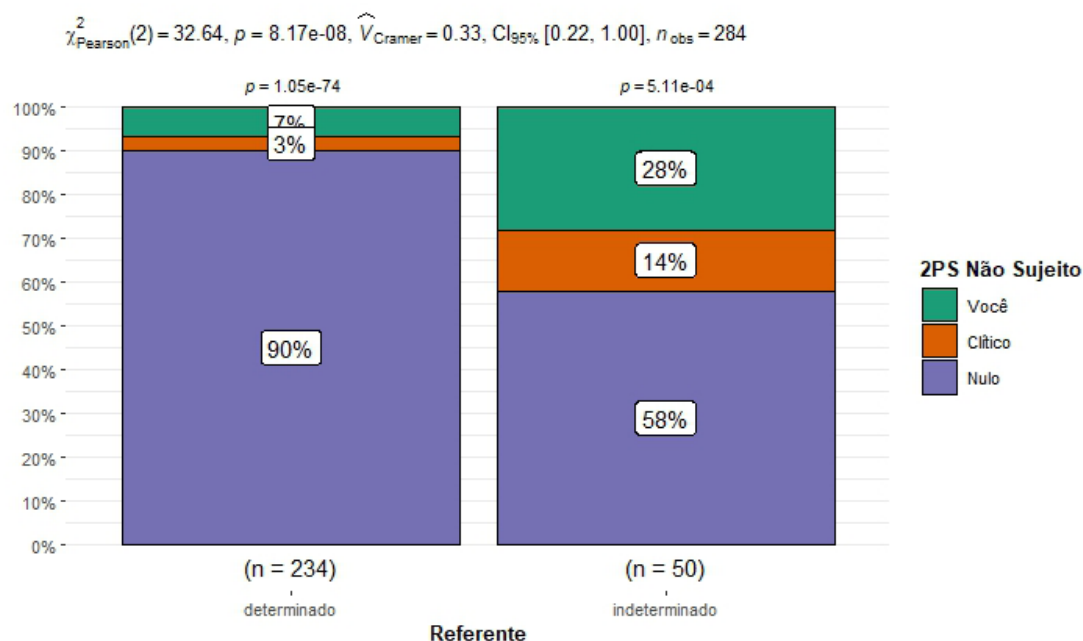


Gráfico 2 – Pronomes de segunda pessoa em posição de não sujeito em função do tipo de referente

Fonte: elaborado pelos autores.

Esses dados confirmam as hipóteses iniciais acerca da preferência das formas de “você” por indeterminação do referente, o que apresenta simetria com as pesquisas dessas formas em posição de sujeito. Contudo, mesmo tendo favorecimento maior nessa posição (28%, $n = 14$), o nulo ainda se mostra prevalente (58%, $n = 29$). Os exemplos (9) e (10) representam as formas de “você” com referência determinada e indeterminada.

(9) “talvez eu não tivesse aqui conversando **com você**” (L75).

(10) “você tem que tê isso todos os dias **com você** - todos os dias” (L58).

Já na relação entre a faixa etária e o uso dos pronomes de 2PS, o Gráfico 3 reporta os resultados encontrados. Como se pode ver, os dados apresentam significância estatística [$\chi^2(4) = 10,90$ $p = 0,03$], apesar da baixa associação entre as variáveis (0,11). Nesse sentido, apesar de haver aumento do uso das formas de “você” na primeira faixa etária ($n = 11$), os resultados convergem para a prevalência do nulo em todas as colunas do gráfico, além de um aumento visível em F3 ($n = 101$).

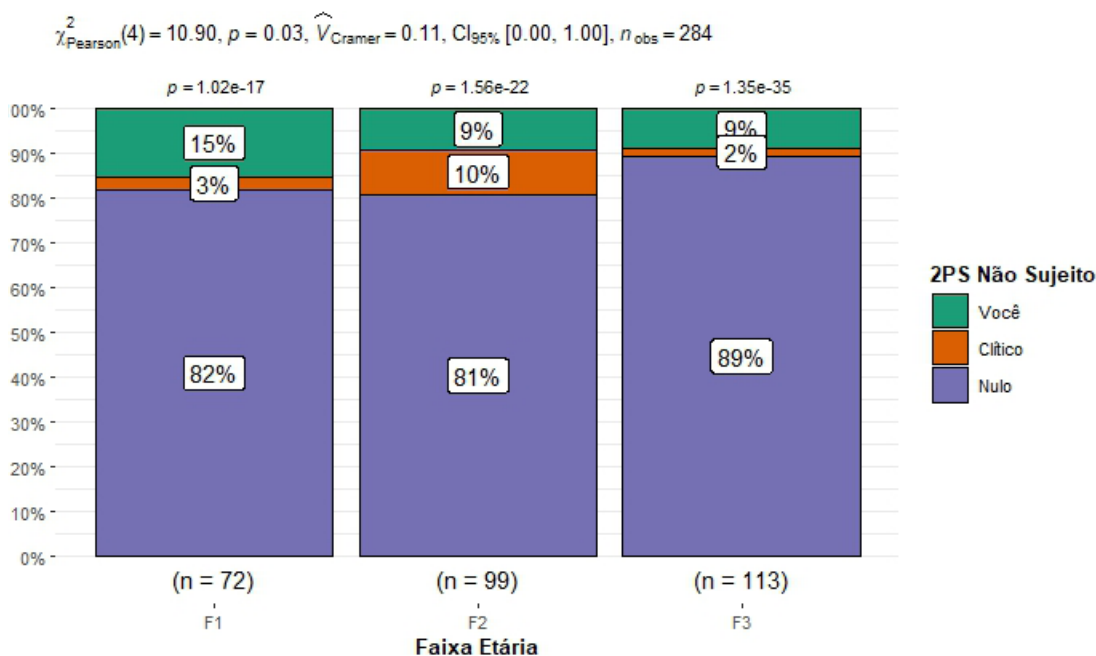


Gráfico 3 – Pronomes de segunda pessoa em posição de não sujeito em função da faixa etária do informante

Fonte: elaborado pelos autores.

Contrariando as hipóteses, os dados mostram que F1 prefere as formas do paradigma “você” ($n = 11$) e F2, os clíticos ($n = 10$), embora haja prevalência de mais de 80% do uso do nulo em todas as faixas. Com essas informações, pode-se inferir, embora com cautela devido à baixa associação, que a variedade sertaneja de Alagoas está em um processo de ampliação do uso das formas de “você”, enquanto diminui o uso do nulo, que passou de 89% ($n = 101$) na F3 para 82% ($n = 59$) na F1. Os exemplos (11), (12) e (13) ilustram o uso das formas do paradigma “você” para cada faixa etária.

(11) “ela tem que sê uma pessoa proativa é: desenrolada - porque é uma profissão que exige x muito **de você**” (L12).

(12) “Ah:: você vai até Piranhas - aí tem o Catamarã e tem aqueles barquinhos - é:: tipo lancha - umas lanchas - aí eles levam **você**” (L61).

(13) “você tem o quê - tem que trabalhá muito com a cabeça pensando **em você** e no ôtro” (L74).

No estudo da relação entre o tipo de relação gramatical e o uso dos pronomes, o Gráfico 4 apresenta os resultados obtidos. Com significância estatística [$\chi^2(4) = 48,15$ $p < 0,001$], nota-se a prevalência do uso do nulo em relação dativa, enquanto o uso das formas de “você” é favorecido nas relações acusativas ($n = 5$) e oblíquas ($n = 12$), assim como o uso dos clíticos na posição acusativa ($n = 3$), o que pode influenciar no valor moderado de associação (0,28). Entretanto, vale destacar a frequência absoluta bem maior dos dados na relação dativa, o que pode se dar pelos pronomes de 2PS estarem ligados, na amostra coletada, a verbos *dicendi*, principalmente os verbos “falar” e “dizer”³, que são bitransitivos.

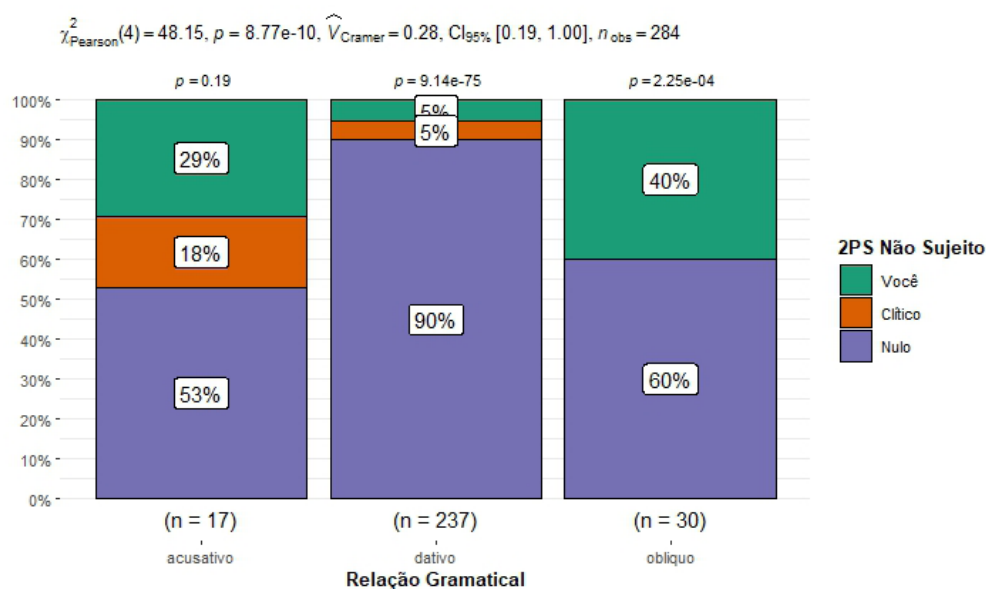


Gráfico 4 – Pronomes de segunda pessoa em posição de não sujeito em função do tipo de relação gramatical

Fonte: elaborado pelos autores.

Tal realidade dos dados confirma as hipóteses acerca da preferência pelo nulo na relação dativa ($n = 213$), assim como pelas formas do paradigma “você” no oblíquo,

³ Tal uso de verbos explica-se pelos informantes referirem-se ao entrevistador para responder às perguntas e, assim, usarem os verbos “falar” e “dizer”. Além disso, como o entrevistador é o único ouvinte, a ele é endereçado a maioria dos pronomes em posição de não sujeito encontrados, logo estes se ligam aos verbos que são direcionados ao entrevistador.

embora as hipóteses não tenham previsto o uso acentuado na posição acusativa, visto que se esperava mais restrição nessa relação gramatical. Também podemos observar que a segunda relação gramatical mais usada por essas formas foi o oblíquo devido à quantidade de ocorrências encontradas, o que corrobora as hipóteses. Assim, percebe-se que há, de fato, um condicionamento dessa variável linguística na comunidade estudada, com exemplos ilustrando o fenômeno em (14), para a relação acusativa, (15), para a dativa, e (16), para a oblíqua.

(14) “primeiramente tem que conversá com: a guia que ela irá orientá **você**” (L12).

(15) “num sei se mãe falô isso **a você**” (L84).

(16) “aquele dinheiro volta: **pra você**” (L87).

Com esses dados, ajustou-se um modelo de regressão logística multinomial (RIPLEY; VENABLES, 2002) com o pacote “nnet” para investigar o poder explicativo das variáveis independentes significativas nos testes de hipótese. Como nível de referência, estabeleceu-se o nulo em vez do clítico (escolha padrão do RStudio pela ordem alfabética) porque o clítico não apresentava dados para a relação gramatical oblíqua, já que há impossibilidade de cliticização nessa posição. A Tabela 2 mostra os resultados do modelo de regressão.

Nessa interação, observa-se a associação positiva do referente indeterminado e negativa da relação gramatical dativa no condicionamento do uso das formas de “você”, com *log-odds* de, respectivamente, 1,24 e -2,09 (o que equivale a probabilidades de 64,33% e 6,06% das formas ocorrerem), isto é, há um favorecimento e um desfavorecimento dessas formas em relação ao intercepto (relação gramatical acusativa, F1 e nulo). Além disso, nota-se uma significância estatística do intercepto do clítico, mostrando a influência da F1, do referente determinado e do acusativo em seu uso; e também se percebe a insignificância da faixa etária e da relação oblíqua⁴ no condicionamento das variantes, o que contraria as hipóteses acerca do oblíquo ser a porta de entrada para as formas de “você” no sistema linguístico. Observa-se também que o modelo explica apenas 20% da variabilidade dos dados, o que abre espaço para novas pesquisas que pensem em outros condicionamentos possíveis para os 80% de variabilidade.

Por último, no Gráfico 5, observa-se uma árvore de inferência condicional (Zeileis; Hothorn, 2015) feita com o pacote “partykit” (Hothorn; Hornik; Zeileis, 2006), que mostra a hierarquização da influência das variáveis no modelo de regressão.

⁴ No caso da variação para o oblíquo no clítico, o grande desfavorecimento (-9.43 *log-odds*) se dá por não haver dados nesse contexto, já que o oblíquo inibe a cliticização.

Tabela 2 – Modelo de regressão multinomial dos pronomes de 2PS não sujeito em função da relação gramatical, do referente e da faixa etária

2PS Não Sujeito					
<i>Preditores</i>	<i>Log-Odds</i>	<i>std. Error</i>	<i>CI</i>	<i>Valor-p</i>	<i>Resposta</i>
(Intercepto)	-2.57	1.07	-4.67 – -0.47	0.016	Clítico
referenteindeterminado	2.24	0.61	1.04 – 3.45	<0.001	Clítico
rel.gramaticaldativo	-1.60	0.81	-3.19 – -0.01	0.048	Clítico
rel.gramaticalobliquo	-9.43	39.56	-87.30 – 68.45	0.812	Clítico
faixa.etariaF2	1.41	0.83	-0.22 – 3.05	0.090	Clítico
faixa.etariaF3	-0.43	1.04	-2.49 – 1.63	0.681	Clítico
(Intercepto)	-0.65	0.69	-2.00 – 0.70	0.342	Você
referenteindeterminado	1.24	0.48	0.30 – 2.18	0.010	Você
rel.gramaticaldativo	-2.09	0.64	-3.35 – -0.82	0.001	Você
rel.gramaticalobliquo	-0.18	0.72	-1.59 – 1.23	0.803	Você
faixa.etariaF2	-0.20	0.54	-1.26 – 0.86	0.708	Você
faixa.etariaF3	-0.54	0.52	-1.58 – 0.49	0.300	Você
Observations			284		
R ² / R ² adjusted			0.209 / 0.202		

Fonte: elaborado pelos autores.

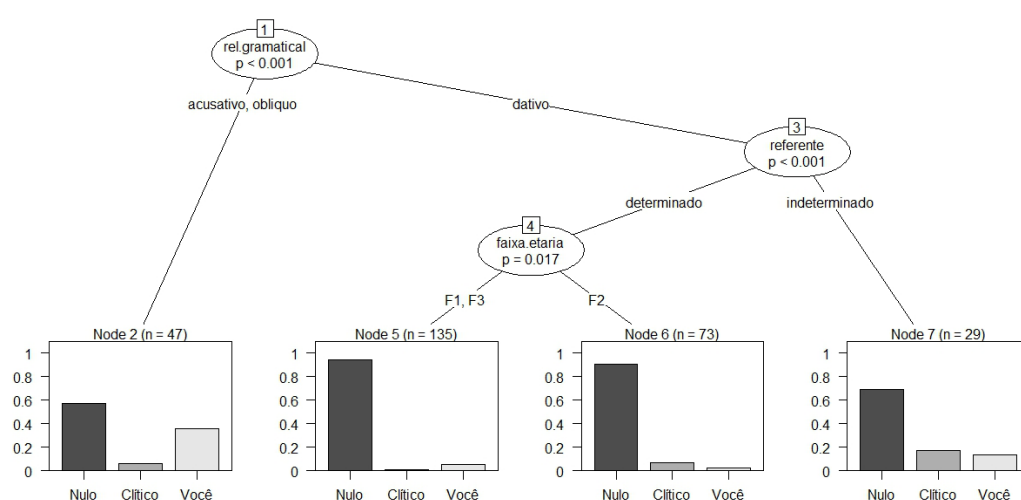


Gráfico 5 – Árvore de inferência condicional

Fonte: elaborado pelos autores.

A primeira bifurcação se dá na relação gramatical, que mostra um nó das relações acusativas e oblíquas com favorecimento das formas de “você”, quase equiparadas à forma nula. Em seguida, percebe-se, no nó do referente, a bipartição do tipo “determinado”, que se divide na faixa etária, na qual F1 e F3 apresentam desfavorecimento do clítico, enquanto F3 o favorece levemente. Por fim, no referente indeterminado, último nó da árvore, observa-se uma diminuição do uso do nulo e favorecimento das formas do “clítico” e de “você”. Logo, percebe-se que os fatores mais importantes no condicionamento dos pronomes de 2PS em posição de não sujeito são linguísticos, embora haja certa influência da pressão social *Faixa Etária*.

Considerações finais

Tomando por base os pressupostos teórico-metodológicos da Sociolinguística Variacionista, a pesquisa aqui reportada objetivou analisar como os falantes naturais do alto sertão alagoano usam os pronomes de segunda pessoa do singular em posição de não sujeito, além dos fatores condicionadores dessa variação. A partir disso, investigou-se um *corpus* falado constituído de 96 entrevistas sociolinguísticas, pertencente ao projeto LUSA.

Os dados encontrados apontam para uma preferência da comunidade pela forma nula, seguida das formas do paradigma “você”, tendo o paradigma “clítico” pouco espaço. Contudo, isso pode se dar pelo tipo de entrevista realizada (DID), que não favorece uma intimidade entre entrevistador e informante, já que a forma “te” teria um traço estilístico/pragmático [+informal], como apontaram as pesquisas da revisão de literatura.

Também se observa que a variação analisada foi condicionada pelos grupos de fatores *relação gramatical*, *tipo de referente* e *faixa etária*, com a não relevância estatística das variáveis *sexo/gênero*, *escolaridade*, *tipo de núcleo* e *paralelismo formal e semântico*. Em relação ao espriamento de “você” nos contextos sintáticos de complementação e adjunção, verifica-se que a sua porta de entrada ocorre nos seguintes contextos: referente indeterminado, falantes mais jovens – F1 (18 a 29 anos) – e relação oblíqua.

É necessário investigar em futuros trabalhos outros tipos de coletas de dados de modo a ter uma compreensão maior do fenômeno na região alagoana. Além disso, cabe a futuros trabalhos investigar em perspectiva diacrônica como o nulo foi ganhando espaço na variedade analisada suplantando as outras formas variantes. Além disso, novos trabalhos poderiam investigar outros tipos de restrições sociais associadas ao fenômeno em busca de descobrir quais aspectos extralinguísticos entram em jogo quando se opta por uma ou outra forma, pois, assim, ter-se-á maior cobertura da variação dos pronomes de 2PS em posição de não sujeito.

Referências

- ALMEIDA, G. S. *Uso Variável dos Pronomes-objeto na Expressão do Dativo e do Acusativo de Segunda Pessoa em Santo Antônio de Jesus – BA*. 2014. 256 f. Tese (Doutorado em Língua e Cultura) do Programa de Pós-Graduação em Língua e Cultura, Instituto de Letras, Universidade Federal da Bahia, Salvador. Bahia: Editora da UFBA, 2014.
- ALMEIDA, G. S. *Uso dos Pronomes-Objetos de Segunda na Fala de Salvador e de Santo Antônio de Jesus. Veredas atemática*, Juiz de Fora, PPG Linguística/UFJF, 2016.
- ARAÚJO, A. S.; BORGES, D. K. V. *Variação no Uso de Pronomes-objeto de Segunda Pessoa na Fala de estudantes Itabaianenses. Paraguaçu*, v. 1, n. 1, 2021.
- BECHARA, E. *Gramática Fácil*. Rio de Janeiro: Nova Fronteira, 2014a.
- BECHARA, E. *Lições de Português pela Análise Sintática*. 19.ed. Rio de Janeiro: Nova Fronteira, 2014b.
- BECHARA, E. *Moderna Gramática Portuguesa*. 39. ed. Rio de Janeiro: Nova Fronteira, 2019.
- CASTILHO, A. T. *Nova Gramática do Português Brasileiro*. São Paulo: Contexto, 2010.
- CRUZ, I. A.; RUMEU, M. C. B. O ‘Você’ em Contextos Sintáticos de Complementação e de Adjunção em Missivas Mineiras (Séc. XX). *Caletroscópio*, v. 4, n. especial, 2016.
- CUNHA, C.; CINTRA, L. *Nova Gramática do Português Contemporâneo*. 7.ed. Rio de Janeiro: Lexikon, 2017.
- DALTO, C. D. L. *Estudo Sociolinguístico dos Pronomes-Objetos de Primeira e de Segunda Pessoas nas Três Capitais do Sul do Brasil*. Paraná: UFPR, 2002. Dissertação (Mestrado em Estudos Linguísticos) do Programa de Pós-graduação em Letras, Setor de Ciências Humanas, Letras e Artes, UFPR, Curitiba, 2002.
- DUARTE, I. *Relações Gramaticais, Esquemas Relacionais e Ordem de Palavras. In: MATEUS, M. H. M.; BRITO, A. M.; DUARTE, I.; FARIA, I. H. Gramática da Língua Portuguesa*. 7.ed. Portugal: Caminho, 2003. p. 277-321
- DUARTE, M. E. L. *O Papel da Sociolinguística na Descrição da Gramática da Escrita Contemporânea. In: MARTINS, M. A.; TAVARES, M. A. (org.). Contribuições da Sociolinguística e da Linguística Histórica para o Ensino de Língua Portuguesa*. Natal: EDUFRN, 2013.
- GALVES, C. *et al.* *Morfossintaxe e Uso dos Pronomes Pessoais na Sincronia e na Diacronia do Português Brasileiro. In: JÚNIOR, L. A. S.; MARTINS, M. A. (org.). Rumos da Linguística Brasileira no Século XXI: historiografia, gramática e ensino*. São Paulo: Blucher, 2016. p. 123-153
- GAMA, D. E. R. S. *O Uso Variável dos Clíticos para Referenciar o Interlocutor. Revista Digital dos Programas de Pós-Graduação do Departamento de Letras e Artes da UEFS, Feira de Santana*, v. 19, n. 2, 2018.

- HOTHORN, T.; HORNIK, K.; ZEILEIS, A. Unbiased Recursive Partitioning: a conditional inference framework. *Journal of Computational and Graphical Statistics*, v. 15, n. 3, 2006. DOI: <https://doi.org/10.1198/106186006X133933>.
- LABOV, W. *Padrões Sociolinguísticos*. São Paulo: Parábola, 2008.
- LOPES, C. R. S. Pronomes pessoais. In: BRANDÃO, S. F.; VIEIRA, S. R. (org.). *Ensino de gramática: descrição e uso*. São Paulo: Contexto, 2007. p. 103-114
- LOPES, C. R. S. O Quadro dos Pronomes Pessoais: descompasso entre pesquisa e ensino. *Matraga*, Rio de Janeiro, v. 19, n. 30, 2012.
- LOPES, C. R. S. et al. A Reorganização do Sistema Pronominal de 2ª Pessoa na História do Português Brasileiro: outras relações gramaticais. In: LOPES, C. R. S. (coord.). *Mudança Sintática das Classes de Palavras: perspectiva funcionalista*. São Paulo: Contexto, 2018. p. 144-185
- MENON, O. P. S. O Sistema Pronominal do Português do Brasil. *Letras*, Curitiba, n. 44, 1995.
- MOLLICA, C. M.; BRAGA, M. L. *Introdução à Sociolinguística: o tratamento da variação*. 4.ed. São Paulo: Contexto, 2020.
- NEVES, M. H. M. *Gramática de Usos do Português*. 2.ed. São Paulo: EDUNESP, 2011.
- PATIL, I. *Visualizations with statistical details: The 'ggstatsplot' approach*. PsyArxiv, 2018. DOI:10.31234/osf.io/p7mku.
- PERINI, M. A. *Gramática Descritiva do Português Brasileiro*. Petrópolis: Vozes, 2016.
- R CORE TEAM. *R: A language and environment for statistical computing*. Áustria: R Foundation for Statistical Computing, 2022.
- RIPLEY, B; VENABLES, W. *Modern Applied Statistics with S*. 4.ed. New York: Springer, 2002.
- ROCHA LIMA, C. H. *Gramática Normativa da Língua Portuguesa*. 49.ed. Rio de Janeiro: José Olympio, 2011.
- RUMEU, M. C. B. A Difusão do *Você* pelos Contextos Sintáticos de Complementação e Adjunção. *Revista Portuguesa de Humanidades | estudos linguísticos*, Portugal, v. 18, n. 1, 2014.
- RUMEU, M. C. B. 'Tu', ou 'Você', 'Te' ou 'Lhe?': a correlação entre as funções de sujeito e complemento verbal de 2ª pessoa. *Linguística*, v. 31, n. 2, 2015.
- RUMEU, M. C. B; OLIVEIRA, T. L. A Expressão da 2ª Pessoa do Singular em Contextos de Complementação e de Adjunção: retratos do encaixamento estrutural e social. *Linguística*, v. 32, n. 2, 2016.
- SCHERRE, M. et al. Variação dos pronomes "tu" e "você". In: MARTINS, M. A.; ABRAÇADO, J. (org.). *Mapeamento sociolinguístico do português brasileiro*. São Paulo: Contexto, 2015, p. 133-172.
- SILVA, S. O. P.; VITÓRIO, E. G. S. L. A. A variação *você* e *cê* no sertão alagoano. *Revista Leitura*, v. 2, p. 122-142, 2017.

VITÓRIO, E. G. S. L. A. A Língua Usada no Sertão Alagoano: constituição da amostra. *In: OLIVEIRA, A. A.; PAULA, A. S. (org.). Interfaces Sociolinguísticas: análises variacionistas em Alagoas e Pernambuco*. Arapiraca: Eduneal, 2020. p. 153-173

WEINREICH, U.; LABOV, W.; HERZOG, M. *Fundamentos empíricos para uma teoria da mudança linguística*. São Paulo: Parábola, 2006.

YACOVENCO, L. C.; SCARDUA, J. R. A Variação Pronominal de Segunda Pessoa: contribuições da sociolinguística para o ensino de língua portuguesa. *Working Papers em Linguística*, Florianópolis, v. 18, n. 2, 2017. DOI: <https://doi.org/10.5007/1984-8420.2017v18n2p171>.

ZEILEIS, A.; HOTHORN, T. Partykit: A Modular Toolkit for Recursive Partytioning in R. *Journal of Machine Learning Research*, v. 16, 2015.

Contribuições dos Autores

Elyne Giselle de Santana Lima Aguiar Vitória – revisão de literatura, análise e descrição dos dados, escrita do artigo. Pedro Henrique Sousa dos Santos – revisão de literatura, análise e descrição dos dados, escrita do artigo.

Conflito de interesse

Os autores declaram a inexistência de conflito de interesse relacionado à elaboração do texto.

Financiamento

Esta pesquisa foi financiada por uma bolsa CNPq de Iniciação Científica de número 127922/2022-9 durante o ciclo do Pibic 2021/2022 da UFAL.