



Apagamento do rótico em coda no Português Santomense (PST): uma análise sociolinguística

Rhotic deletion in coda in Santomean Portuguese (PST): an sociolinguistics analysis

Nancy Mendes Torres Vieira

Universidade de São Paulo (USP), São Paulo, São Paulo / Brasil

nancy.vieira@usp.br

<http://orcid.org/0000-0002-7291-9759>

Amanda Macedo Balduino

Universidade de São Paulo (USP), São Paulo, São Paulo / Brasil

amanda.m_b@hotmail.com

<http://orcid.org/0000-0002-1062-973X>

Resumo: O objetivo deste artigo é analisar o apagamento do rótico em coda no português de São Tomé (PST), uma variedade do português falada em São Tomé e Príncipe. Para análise desse processo, nos pautamos na Sociolinguística Variacionista e a análise quantitativa foi realizada nos softwares RStudio e Rbrul. Consideramos variáveis sociais (Sexo, Faixa Etária e Grau de Escolaridade) e linguísticas (Classe Gramatical, Tonicidade, Posição do Segmento e Contexto Fonológico Precedente) que poderiam favorecer ou desfavorecer o fenômeno. Examinando as 1523 ocorrências de rótico em coda, verificamos o apagamento do rótico em 56,53% dos dados. A variável mais significativa foi a Posição do Segmento com o maior índice de apagamento constatado quando o rótico estava em posição final (P.R. .71). A amostra foi, portanto, dividida em duas amostras menores analisadas separadamente: 712 ocorrências do rótico em coda medial e 811 em coda final. Na análise do fenômeno em coda medial, o índice de apagamento foi de 32,6% e os maiores índices de apagamento foram verificados entre os informantes do ensino fundamental (P.R. .62) e do sexo feminino (P.R. .59). Já em coda final, o índice de apagamentos foi de 77,5% e os falantes com ensino fundamental (P.R. .75) e mais jovens (P.R. .76) foram mais propensos ao apagamento, ademais, a classe verbal favoreceu o apagamento (P.R. .62). Os resultados são cojetados com trabalhos

prévios (BOUCHARD, 2017; BRANDÃO, 2018; BRANDÃO; DE PAULA, 2018; BRANDÃO *et al.*, 2017) e com resultados de estudos sobre variedades do PB e do PE.

Palavras-chave: português santomense; apagamento do rótico em coda; variáveis sociais e linguísticas; contato linguístico.

Abstract: The goal of this paper is to analyze rhotic deletion in coda in Santomean Portuguese (PST), a Portuguese variety spoken in Sao Tome and Principe. This process was analyzed based on Variationist Sociolinguistics. Quantitative analyses have been carried out by softwares RStudio and Rbrul. Thus, we consider social (sex, age and education level) and linguistic (grammatical class, syllable stress, position in morphological word and quality of preceding vowels) variables that might favor or not the phenomenon. By examining 1523 rhotic occurrences in coda, we verified r-deletion in 56,53% of the data. The most significant variable is Segmental Position with the highest rate of r-deletion in word-final syllables (P. R. 71). Then, the sample was divided into two samples separately analyzed: 712 occurrences in word-medial coda and 811 occurrences in word-final coda. In word-medial coda, 32,6% of rhotics are deleted. The highest rates of loss are associated to elementary school (P.R. .62) and women (P.R. .59). In word-final coda, 77.5% of rhotics are deleted, Speakers with elementary school (P.R. .75) and younger speakers (P.R. .76) are more likely to delete the rhotic in coda. Considering Grammatical Class, r-deletion is favored in verbs (P.R. .62). These outcomes are compared with previous studies (BOUCHARD, 2017; BRANDÃO, 2018; BRANDÃO; DE PAULA, 2018; BRANDÃO *et al.*, 2017) and with studies on PB and PE.

Keywords: Santomean Portuguese; deletion of rhotic in coda; social and linguistic variables; linguistic contact.

Recebido em 18 de novembro de 2020

Aceito em 18 de janeiro de 2021

1 Introdução

Neste artigo, examinamos o apagamento do rótico em coda silábica no português de São Tomé (PST). Mediante a Sociolinguística Variacionista (LABOV, 1972), visamos, assim, analisar o fenômeno de apagamento segmental em coda, amplamente reportado para variedades como o português brasileiro (PB) e o português europeu (PE), em uma variedade da língua portuguesa empregada em um contexto multilíngue. O presente estudo surgiu a partir do trabalho de Vieira e Balduino (2020), que discutiu o preenchimento da coda no português santomense (PST),

avaliando a possibilidade de apagamento nessa posição como resultado do contato dessa variedade do português com o santome, uma língua crioula desenvolvida e falada em São Tomé (ARAÚJO; HAGEMEIJER, 2013; BANDEIRA, 2017; FERRAZ, 1979).

O estudo de Vieira e Balduino (2020) contabilizou os seguintes índices de apagamento: 61,53% (355/577) para o rótico, 24,21% (77/318) para a lateral e 5,27% (29/550) para a sibilante, concluindo, entre outras coisas, que uma possível convergência linguística entre o santome e o PST não é suficiente para explicar os diferentes percentuais de apagamento em coda na variedade de português falada em São Tomé (VIEIRA; BALDUINO, 2020, p. 29). Para as autoras, fatos relacionados ao contato são importantes para análise de fenômenos fonológicos no PST, porém devem ser cotejados em conjunto com aspectos estruturais das línguas, bem como com o espaço sociolinguístico onde a língua é falada, abarcando, portanto, fatores sociais relacionados ou não ao contato linguístico. O estudo de Vieira e Balduino (2020), embora ofereça uma descrição pioneira do processo de apagamento nas codas /r, l, S/, indicando ser este um fenômeno de implementação ampla no PST, não abarca a questão mediante modelagem estatística dos dados nem sob a perspectiva teórica da Sociolinguística Variacionista. Dessa forma, visando contribuir com a análise do apagamento da coda a partir de uma nova abordagem metodológica e focando no rótico como segmento-alvo, delimitamos como objetivos desse estudo: (i) analisar o apagamento do rótico no PST – segmento que apresentou o maior índice de apagamento no trabalho de Vieira e Balduino (2020) – a partir de uma amostra maior e à luz da Sociolinguística Variacionista, verificando os fatores sociais e linguísticos que influenciam esse fenômeno; e (ii) cotejar nossos resultados com trabalhos prévios que, utilizando a mesma metodologia, indicam ser o apagamento em coda um fenômeno recorrente no PST (BOUCHARD, 2017; BRANDÃO, 2018; BRANDÃO; DE PAULA, 2018; BRANDÃO *et al.*, 2017) e com dados de estudos sobre o mesmo processo em variedades do PB e do PE.

O artigo está organizado da seguinte forma: na seção 2, discutimos, brevemente, a consolidação do português como língua materna da população santomense; na seção 3, apresentamos o fenômeno examinado; na seção 4, definimos os procedimentos de exame adotados; na seção 5, expomos os resultados da análise dos dados e, por fim, tecemos as considerações finais, na seção 6.

2 Contexto linguístico de São Tomé

A República de São Tomé e Príncipe (STP) é um país da Costa Oeste Africana localizado no Golfo da Guiné. Desde 1975, o português é a única língua oficial do arquipélago, entretanto, línguas autóctones como o santome (ISO 639-3: CRI), o angolar (ISO 639-3: AOA) e o lung'le (ISO 639-3: PRI), além do kabuverdianu (ISO 639-3: KEA), uma língua transplantada de Cabo Verde a STP, também são faladas pelos seus habitantes (AGOSTINHO, 2015; ARAUJO, 2020; BANDEIRA, 2017; HAGEMEIJER, 2018).

De acordo com Gonçalves (2010), Christofolletti (2013), Balduino (2018), entre outros autores, a convivência da língua portuguesa com as línguas autóctones, em STP, cria um contexto ecolinguístico complexo e desigual em que questões de standardização linguística interferem no emprego linguístico. A língua portuguesa, em geral, exerce efeitos glotocidas sobre as demais línguas do arquipélago, na medida em que é adquirida como língua materna (L1) em detrimento da aquisição das demais línguas faladas em STP.

Distintamente de outros países africanos em que o português configura uma das línguas oficiais, porém não é a língua majoritariamente falada pela população, como é o caso de Angola, Moçambique, Guiné-Bissau e Cabo Verde, em STP, o português é a língua materna de, cerca de, 98% da população de acordo com o último censo publicado (INE, 2012). A difusão do português enquanto língua materna deve-se, sobretudo, à convergência de fatores como a urbanização, democratização de ensino, dado que a alfabetização era e ainda é realizada em português, bem como às próprias dinâmicas sociolinguísticas do país, as quais estão circunscritas numa conjuntura multilíngue (ARAUJO, 2020; GONÇALVES; HAGEMEIJER, 2015).

Conforme Araujo (2020), as elites locais desempenharam importante papel para a eleição do português como língua oficial bem como para sua difusão enquanto L1. Assim como nos demais territórios africanos que passavam por um processo de descolonização em relação a Portugal, o português foi eleito como língua oficial de STP não somente por corresponder a uma língua com tradição escrita, mas, também, como uma pretensa tentativa de diminuir conflitos étnicos locais. Para Araujo (2020), essa “neutralização política”, demarcada pela escolha do português como língua oficial, não pode deixar de ser avaliada em cotejo

com o fato de que esta já era empregada como língua de um grupo social, a elite santomense. Bouchard (2017) indica que a elite e seus descendentes, ainda que marginalizados pela elite portuguesa pré-independência, já tinham acesso à educação em Portugal e empregavam o português. Assim, no período pós-independência, o português foi preservado por tal grupo social, sendo promovido por instituições educacionais e midiáticas, ao passo que não foram criadas políticas linguísticas dedicadas à transmissão e preservação das línguas crioulas do arquipélago (ARAÚJO, 2020), fatos que contribuíram para sua marginalização.

Ainda que o português configure a língua mais usada e transmitida em STP, é possível observar divergências entre as variedades vernáculas de português do arquipélago e a norma de prestígio adotada por seus falantes e disseminada pela escolarização (BALDUINO, 2018; BOUCHARD, 2017). A variedade de português que goza de maior status social em STP, língua-alvo da população de STP, ainda é a variedade europeia (BALDUINO, 2018).

Em STP, circulam diferentes variedades vernáculas da língua portuguesa. Além de variedades urbanas, como o português santomense (PST) e o português principense (PP) (BALDUINO, 2019; SANTIAGO, 2019; SANTIAGO; AGOSTINHO, 2020), há registros de variedades como o português de Almoxarife (PA), empregado na costa da ilha de São Tomé, e do português dos Tongas (PT), falado na comunidade da roça Monte Café (BAXTER, 2018; FIGUEIREDO 2010; ROUGÉ, 1992). Ademais, outras variedades não registradas, como o português angular, falado na região dos Angolares, e o português falado pelos descendentes cabo-verdianos, também são observadas.

Essas variedades possuem estruturas e processos linguísticos singulares que as diferenciam da variedade europeia e as consolidam enquanto variedades da língua portuguesa. Tais processos linguísticos têm respaldo em fatores sociolinguísticos, entre os quais pode estar a influência das línguas autóctones. Isto posto, neste artigo, propomos uma análise do processo de apagamento do rótico em coda no PST, variedade vernácula da cidade de São Tomé. Para tanto, considerando que as dinâmicas sociais do país foram fundamentais para ascensão do português como L1 da população santomense, investigaremos possíveis variáveis linguísticas e sociais, que consolidam o espaço sociolinguístico no qual o PST está inserido, e podem estar correlacionadas com esse fenômeno.

3 Apagamento do rótico em coda silábica

Neste artigo, assume-se que a sílaba apresenta uma estrutura hierárquica, sendo organizada a partir de constituintes internos como: o *onset* (O) e a rima (R), subdividida em núcleo (N) e coda (C) (SELKIRK, 1982). Tendo como foco de análise a coda, observamos que neste constituinte são licenciadas consoantes como sibilantes /S/, lateral /l/, nasal /N/ e róticos no PST, (BALDUINO, 2019; BALDUINO; VIEIRA, 2020; VIEIRA; BALDUINO, 2020), assim como observado no português brasileiro (PB) e no europeu (PE) (CÂMARA JR., 1970; MATEUS; D'ANDRADE, 2000).

De acordo com diferentes trabalhos que analisam fenômenos cuja sílaba é o domínio, no português, a coda é propícia à variação e a apagamentos, os quais priorizam, especialmente no PB, a estrutura CV (ABAURRE; SÂNDALO, 2003; BÍ SOL, 1999; BRANDÃO; MOTA; CUNHA, 2003; CALLOU; LEITE; MORAES, 1994; CALLOU; LEITE; MORAES, 2002; CÂMARA JR., 1970; MATEUS; RODRIGUES, 2003; QUEDNAU, 1993).

Em geral, o apagamento de codas pode ocorrer tanto em PB quanto em PE, sendo, entretanto, mais recorrente em PB (BRANDÃO; MOTA; CUNHA, 2003; CALLOU; SERRA, 2012; OLIVEIRA, 2018; RODRIGUES, 2012; SERRA; CALLOU, 2015), o que é fundamentado pelo fato de o PE ser uma modalidade de reforço consonântico, enquanto o PB tende a reforçar seu quadro vocálico (BRANDÃO; MOTA; CUNHA, 2003, p. 14).

O rótico apresenta não somente grande variabilidade, mas possui índices expressivos de apagamento em diversas variedades do PB (ALVES, 2015; BRANDÃO; MOTA; CUNHA, 2003; CALLOU; LEITE; MORAES, 1994; CALLOU; SERRA, 2012; CALLOU; SERRA; CUNHA, 2015; MONGUILLHOT, 1997; OLIVEIRA, 2018; OUSHIRO; MENDES, 2014; SERRA; CALLOU, 2015), do PST (BOUCHARD, 2017; BRANDÃO, 2018; BRANDÃO; DE PAULA, 2018; BRANDÃO *et al.* 2017; VIEIRA; BALDUINO, 2020) e, ainda, em outras variedades africanas da língua portuguesa, como o português de Moçambique (BRANDÃO, 2018)

Agostinho (2016), Bouchard (2017), Brandão *et al.* (2017), Brandão e De Paula (2018) e Brandão (2018) chamam atenção para o fato de haver, no PST e no PP, alto grau de instabilidade na produção

dos róticos, sendo esse um segmento alvo de processos fonológicos como o apagamento que, por sua vez, é influenciado por variáveis linguísticas e sociais. De forma semelhante a variedades como o PB e o PE, o PST também apresenta diferentes realizações do rótico. Entretanto, distintamente de tais variedades, o rótico pode perder valor distintivo entre vogais, como indicado pelo par mínimo caro ['ka.ro] ~ ['ka.ʁo], e carro ['ka.ro] ~ ['ka.ʁo], em que o R-forte e o R-fraco alternam, inclusive, em diferentes produções de um mesmo falante. Nos demais contextos de ocorrência, essa alternância é também sustentada: como em *onset* no início de palavra (rato ['ra.tu] ~ ['ʁa.tu]), no segundo elemento de um *onset* complexo (prato ['pra.tu] ~ ['pʁa.tu]), e nas codas mediais e finais (carta ['kar.tɐ] ~ ['kaʁ.tɐ] e mar ['mar] ~ ['maʁ], respectivamente) (BALDUINO, em preparação).

Para Bouchard (2017), o uso do R-forte no PST, seja como C1 ou C2, seja em contexto intervocálico ou não, é influenciado pela variável faixa etária: o uso de [ʁ] pelas gerações mais jovens de São Tomé seria, assim, um indicativo de mudança linguística em curso, em que [r] está sendo substituído por [ʁ]. Essa hipótese é reforçada por Agostinho, Soares e Mendes (2020)¹ que discutem a possibilidade de fusão de quase-fonemas, isto é, de segmentos que apresentam relações fonológicas intermediárias, variando no grau de previsibilidade e contraste. Conforme o experimento conduzido pelos autores, há evidências que sugerem fusão e perda de contraste dos róticos no Português do Príncipe, variedade também falada em STP. Isso ocorreria em decorrência do seu status quase-fonêmico, assim como por sua baixa carga funcional no sistema e pelo contato linguístico com as línguas crioulas faladas na região (AGOSTINHO; SOARES; MENDES, 2020).

As propostas de Bouchard (2017) e de Agostinho, Soares e Mendes (2020) demonstram que o estabelecimento do fonema rótico no inventário fonológico do PST não é uma tarefa simples. Neste estudo, analisaremos o apagamento do rótico em coda concebendo a existência de

¹ “Merging of Quasi-Phonemes in Contact Situations: Evidence from Rhotics in Principense Portuguese”. Trabalho apresentado por Ana Livia Agostinho, Eduardo Soares, e Maiara Mendes no Annual Meeting on Phonology, em 2020 Santa Cruz: University of California, 2020.

duas realizações concorrentes no PST: [r] e [ʁ],² os quais serão referidos como rótico, posto que o estabelecimento de uma forma fonológica para tais fones corresponde a uma discussão que foge ao escopo desse trabalho. Não adotaremos, assim, a representação /R/ para róticos em coda, uma vez que esta pressupõe a neutralização de um fonema numa determinada posição, fato que não condiz ao comportamento do rótico no PST.³

4 Metodologia

O estudo baseia-se nos pressupostos da Teoria da Variação e Mudança, também denominada Sociolinguística Variacionista, desenvolvida por Weinreich, Labov e Herzog (1968) e Labov (1972, 1994, 2001). A Sociolinguística é uma teoria linguística social que tem como objeto de estudo a língua, como é usada na vida cotidiana de uma comunidade, considerando os fatores sociais que se correlacionam com ela (LABOV, 1972).

A análise do apagamento do rótico no PST foi realizada com base em dados de um *corpus* coletado em São Tomé, o qual é constituído por entrevistas de fala espontânea gravadas em 2016⁴ e 2019⁵. Durante a coleta dos dados, seguimos os seguintes procedimentos éticos, exigidos na realização de pesquisas com seres humanos: (i) os participantes da pesquisa eram maiores de 18 anos e assinaram um Termo de Consentimento Livre e Esclarecido (TCLE), contendo informações gerais e esclarecimentos sobre a pesquisa, no qual concordaram, de

² É preciso ressaltar, no entanto, que os róticos, em coda, apresentam grande variabilidade no PST, sendo observáveis, ainda, realizações como [r], [h], [ʁ], ou mesmo uma produção vocalizada [w] ou lateralizada [ʎ].

³ A neutralização é um processo natural pelo qual dois ou mais fonemas, opostos em um determinado contexto, deixam de apresentar oposição em outro ambiente. No PB, o rótico é neutralizado em coda, posição em que pode ser produzido de diferentes formas: [r], [ʁ], [h], [ɹ], [x], entre outros. Já em meio de palavra, nesta variedade, o r-forte e o r-fraco constituem fonemas distintos e geram oposição de pares mínimos: caro ['ka. rɔ] e carro ['ka. rɔ] ~ ['ka. ʁɔ]. No PST, como discutido ao longo da seção 3, isso não ocorre, visto que a oposição entre os róticos não é verificada: **caro** ['ka. rɔ] ~ ['ka. ʁɔ], e carro ['ka. rɔ] ~ ['ka. ʁɔ].

⁴ Trabalho de campo de Ana Livia Agostinho e Amanda Macedo Balduino, não publicado.

⁵ Trabalho de campo de Amanda Macedo Balduino, não publicado.

forma consciente, em participar do estudo; (ii) antes da gravação das entrevistas, os participantes da pesquisa foram informados de que seriam gravados e, durante as entrevistas, demonstraram estar cientes de tal fato, manifestando seu consentimento em participar do estudo; (iii) os nomes dos entrevistados foram omitidos na divulgação de dados do estudo.

Nesta análise, trabalhamos com doze entrevistas, cada uma com duração de, aproximadamente, 60 minutos, porém descartamos os 15 primeiros minutos, a fim de examinarmos uma produção mais próxima à fala vernacular dos informantes. Todos os informantes eram naturais da ilha de São Tomé, tendo o português como língua materna.

O *corpus* é composto por ocorrências de palavras com rótico em coda, sendo que, para cada entrevista, não foram consideradas mais do que seis realizações para uma mesma palavra, ou mais do que três realizações de um mesmo verbo no infinitivo. Esse método, por um lado, registra as possíveis realizações de um mesmo item lexical, já que um mesmo falante pode pronunciar uma mesma palavra de formas distintas. Por outro lado, ao restringirmos o número de ocorrências de cada item lexical, reduzimos a possibilidade de que o *corpus* fique enviesado, posto que tal método limita a influência de palavras de alta frequência que podem não representar o fenômeno de maneira efetiva na variedade em questão.⁶

Inicialmente foram contabilizadas 1564 ocorrências, então excluímos os casos de ressilabificação, em que o rótico passa a ocupar a posição de *onset* na sílaba seguinte, e esse número foi reduzido a 1523 ocorrências constituídas por 534 palavras.

Essas ocorrências foram submetidas a uma análise acústica, de modo a verificarmos o apagamento da coda nas palavras analisadas. Para tanto, utilizamos o *Praat*,⁷ que possibilitou verificar a realização ou ausência do segmento na produção das palavras pelos informantes. O apagamento foi considerado quando não havia qualquer forma espectral

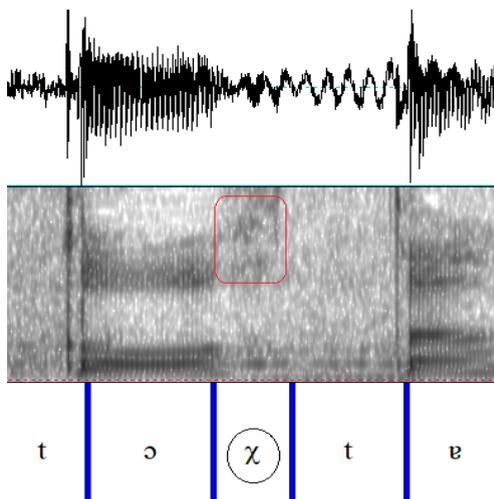
⁶ Verbos no infinitivo, por exemplo, correspondem a itens em que a recorrência de apagamentos do rótico poderia ser numerosa, já que a estrutura dessa classe gramatical constitui contexto propício para o fenômeno. Daí a relevância de também serem observados itens não-verbais e limitarmos a ocorrência de verbos.

⁷ *Praat* é o software utilizado para análise e síntese de fala, pelo qual é possível acessar informações acústicas do segmento, como a duração, o formato de onda, o espectrograma e, no caso deste estudo, sua ausência (BOERSMA; WEENINK, 2020).

que pudesse ser atribuída ao segmento alvo, cuja não realização já havia sido estabelecida, previamente, mediante uma análise de oitiva. Operamos, assim, com um exame espectral categórico, o qual não leva em consideração formas fonéticas gradientes (MENESES, 2012).

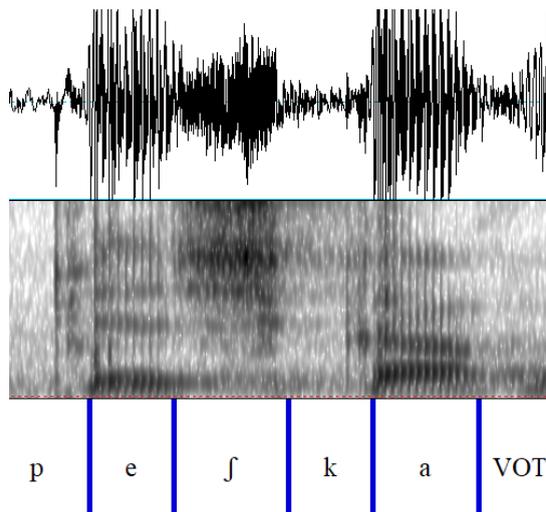
Nas Figuras 1 e 2, observamos exemplos de espectrogramas que ilustram a realização ou não de um segmento consonantal em coda. Enquanto na Figura 1 é possível verificarmos a realização da fricativa [χ], em coda medial, na Figura 2, o item lexical *pescar* é produzido sem o rótico, sendo a vogal [a] seguida pelo VOT da consoante oclusiva seguinte à palavra, porém não sendo identificado o rótico. A Figura 2 evidencia os casos nos quais o apagamento foi considerado para a análise.

FIGURA 1 – Espectrograma do item lexical *torta* com realização do rótico em coda medial



Fonte: elaboração própria

FIGURA 2 – Espectrograma do item lexical *pescar* sem realização do rótico em coda final



Fonte: elaboração própria

Feitos os exames espectrais que identificaram os casos de apagamento, partimos para a análise dos dados, nomeadamente, uma análise de regra variável⁸ – em que a regra é o apagamento do rótico em posição de coda – e o tratamento quantitativo dos dados foi realizado no software RStudio.⁹ A modelagem estatística dos dados foi feita pelo

⁸ Análise de regra variável é uma forma de dar conta da variação estruturada, governada por regras, no uso da língua, que consiste num tipo de análise multivariada com o objetivo de separar, quantificar e testar a significância dos efeitos de fatores contextuais em uma variável linguística, aqui o rótico em coda que pode ou não ser apagado (GUY; ZILLES, 2007).

⁹ O RStudio é uma interface funcional e mais amigável para o R que, por sua vez, consiste não apenas num software, mas também numa linguagem de programação voltada para a análise de dados, que pode ser utilizada para realizar computações estatísticas e gráficas, compilar *corpora*, produzir listas de frequências, entre outras diversas tarefas. O RStudio é o principal ambiente de desenvolvimento integrado para R que disponibiliza ferramentas adicionais diretamente na interface gráfica, como a visualização dos *scripts* abertos recentemente, o histórico de linhas de comando executadas, a lista de pacotes instalados, entre outras (OUSHIRO, 2014, p.134-136).

Rbrul,¹⁰ um software que roda, a partir de um código, no RStudio. O modelo estatístico empregado foi o de regressão logística, no qual se verifica o efeito simultâneo de múltiplas variáveis independentes ou previsoras na aplicação da regra variável, calculando o chamado *peso relativo* para cada um dos fatores componentes das variáveis independentes.

O efeito dos fatores das variáveis independentes são valores probabilísticos que estão numa escala que vai de 0 a 1 em que 0 representa uma chance nula de algo ocorrer (0%), 1 representa certeza de que vai ocorrer (100%) e 0,5 é o ponto neutro, já que equivale a 50% de certeza de que algo aconteça. Dessa forma, um fator como “verbo” (da variável classe gramatical) com um *peso relativo* igual a 0,75 significaria que há uma chance de 75% de que a regra variável seja aplicada quando a palavra analisada é um verbo. Sintetizando, um peso relativo (P.R.) com um valor superior a 0,5 indica que o fator favorece a aplicação da regra, ao passo que um valor menor que 0,5 aponta que o fator a desfavorece. Já um valor igual a 0,5, por sua vez, indica que o fator não tem efeito na aplicação da regra.

O software também calcula o *input*, que representa o nível geral de uso de determinado valor da variável dependente, nesse caso, o apagamento de um segmento consonantal em coda. Tal valor deve se aproximar da frequência de aplicação da regra, calculada para a amostra total e quando isso não ocorre, é um indicativo de que a distribuição dos dados, através dos fatores analisados, não é equilibrada. Os P.R. são calculados em relação a esse nível geral. (GUY; ZILLES, 2007; OUSHIRO, 2017). Além disso, o software fornece o nível de significância do modelo, também chamado *valor-p* que é a probabilidade de se observar determinado resultado, por acaso, em caso de a hipótese nula ser verdadeira. O que nos remete a dois importantes conceitos da estatística inferencial: *hipótese nula* (H_0) e *hipótese alternativa* (H_1). Esta última é a hipótese que está sendo testada, como, por exemplo, a afirmação de

¹⁰ O Rbrul é um software gratuito e idealizado por Daniel Ezra Johnson, software que realiza não apenas todas as funções que as versões Goldvarb e Varbrul realizam (regressão múltipla, tabulação cruzada, *step up/step down*), mas também outras funções como: rodar variáveis contínuas como variáveis independentes; variáveis contínuas como variáveis dependentes e, ainda, da conta de modelos mistos. Além disso, o Rbrul estabelece uma interface com as capacidades gráficas do R (GOMES, 2012). O programa e o manual de uso estão disponíveis na página <http://www.danielezrajohnson.com/rbrul.html>.

que há uma relação entre duas variáveis, enquanto a H_0 , normalmente, é formulada como a negação da H_1 , afirmando que não há relação entre as variáveis e que a distribuição dos dados observada resulta de uma flutuação aleatória e/ou erro de amostragem.

Se o valor p for muito baixo, rejeita-se a hipótese nula e a distribuição dos dados observada é considerada estatisticamente significativa, o que significa que a relação ou o efeito que está sendo testado são verdadeiros, já que, a probabilidade de ocorrerem por acaso é muito pequena, assim, quanto menor o valor de p mais significativo é o modelo. A comunidade científica costuma usar o limite de 0,05 (ou 5%) para considerar algo como muito pouco provável para acontecer ao acaso (GUY; ZILLES, 2007; OUSHIRO, 2017).

As variáveis sociais consideradas foram: Sexo, Faixa Etária e Grau de Escolaridade dos informantes. Trabalhamos com um mesmo número de homens e mulheres (seis homens e seis mulheres) a fim de verificar se há diferenças relacionadas ao Sexo, uma vez que estudos têm mostrado que a variação linguística se correlaciona com esta variável (CHESHIRE, 2004). Quanto à Faixa Etária, os informantes foram organizados em três níveis: o primeiro até 20 anos, o segundo entre 21 e 40 anos e o terceiro, acima de 40 anos. O Grau de Escolaridade é um indicador de *status* socioeconômico que, segundo estudos sociolinguísticos, está associado à utilização de traços diferentes da língua (LABOV, 1972) e está dividido da seguinte forma: ensino fundamental (da 4^a à 9^a classe), ensino médio (10^a à 12^a classe) e ensino superior (graduação e pós-graduação). A Tabela 1 mostra uma estratificação dos informantes conforme a faixa etária, o sexo e o grau de escolaridade.

QUADRO 1 – Estratificação dos informantes

Faixa Etária	Sexo		Grau de Escolaridade		
	Feminino	Masculino	Ensino Fundamental	Ensino Médio	Ensino Superior
18-20	2	2	0	4	0
21-40	3	2	2	1	2
40-52	1	2	1	0	2
Totais	6	6	3	5	4

Fonte: elaboração própria

No que se refere às variáveis linguísticas, analisamos: a Classe Gramatical da palavra (verbo e não-verbo), a Tonicidade da sílaba e a Posição do Segmento na palavra (final e medial), além do Contexto Fonológico Precedente quanto ao traço de Altura [-alto] (a, ε, o) e [+alto] (i, e, o, u) e os traços [-anterior] (a, o, u) e [+anterior] (i, e, ε).

Do total de 1523 ocorrências analisadas, 779 estão contidas em itens verbais e 744 em itens não verbais. Quanto à posição do segmento, trabalhamos com 811 ocorrências em coda final e 712 em coda medial. Considerando que estudos que analisam o apagamento do rótico no PB têm apresentando índices quase categóricos de apagamento desse segmento em coda final de itens verbais (CALLOU, 1987 apud OLIVEIRA, 2018, p. 26); CALLOU; LEITE; MORAES, 2002; CALLOU; SERRA, 2012; CALLOU; SERRA; CUNHA, 2015; OLIVEIRA, 2018; OUSHIRO; MENDES, 2014), realizamos o cruzamento das variáveis: Posição do Segmento e Classe Gramatical da palavra, a fim de verificar a diferença entre os índices de apagamento em itens verbais e não-verbais em coda medial e final.

Após a análise inicial, a amostra, de 1523 ocorrências, foi dividida em duas amostras menores: 712 ocorrências do rótico em posição medial (coda interna) e 811 em posição final (coda externa). Feito isso, analisamos as duas amostras, separadamente, verificando a relevância das mesmas variáveis sociais e linguísticas, exceto a variável Posição do Segmento na palavra (final e medial), que deixa de ser pertinente – dado que na primeira amostra há apenas dados do rótico em coda medial e na segunda de coda final.

5 Análise e Resultados

Nesta seção, analisamos o apagamento do rótico, verificando as variáveis sociais e linguísticas que influenciam o fenômeno sem deixar de considerar questões relacionadas ao contato linguístico que podem estar relacionadas ao índice de apagamento. Em seguida, comparamos os resultados com estudos sobre variedades do PB e do PE.

Inicialmente, foram analisadas 1523 ocorrências, controlando oito variáveis, sendo: (i) três sociais: Sexo, Faixa Etária e Grau de Escolaridade dos informantes e (ii) cinco linguísticas: a Posição do Segmento, a Classe Gramatical da palavra, a Tonicidade da sílaba e o Contexto Fonológico Precedente que foi testado como duas variáveis: a primeira considerando o traço de altura [-alto] e [+alto] e a segunda, os traços [-anterior] e

[+anterior]. O apagamento foi constatado em 56,53% (861/1523) das ocorrências evidenciando a alta frequência percentual do fenômeno, no PST, e das oito variáveis testadas, foram selecionadas como relevantes, para a aplicação da regra de apagamento do rótico: Posição do Segmento, Escolaridade, Faixa Etária, Classe Gramatical e Sexo dos informantes conforme a Tabela 1. Quanto às variáveis Tonicidade, Contexto Fonológico Precedente quanto ao traço de altura [-alto] [+alto] e quanto aos traços [-anterior] [+anterior], foram excluídas – apontando que essas variáveis não estavam relacionadas, de modo significativo, com o apagamento do rótico.

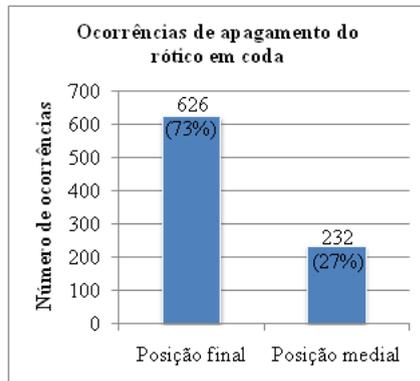
TABELA 1 – Variáveis atuantes no apagamento do rótico em coda no PST

Variável	Fatores	Apl./N	% R-Ø	P.R.
Posição	Final	629/811	77,5	.71
	Medial	232/712	32,6	.29
		861/1523		Range .42
Escolaridade	Fundamental	234/351	66,6	.69
	Médio	403/654	61,6	.40
	Superior	224/528	43,2	.39
		861/1523		Range .30
Faixa Etária	18-20 anos	351/531	66,1	.68
	21-40 anos	335/618	54,2	.42
	41-52 anos	175/374	46,7	.39
		861/1523		Range .29
Classe Gramatical	Verbos	575/779	73,8	.57
	Não-verbos	286/744	38,4	.43
		861/1523		Range .14
Sexo	Feminino	511/821	62,2	.55
	Masculino	350/702	49,5	.45
		861/1523		Range .10
<i>Input: .59</i>				<i>p < 0,0001</i>

Fonte: elaboração própria.

A Posição do Segmento foi a variável mais relevante para a aplicação da regra de apagamento com um *range*¹¹ de .42. O maior índice de apagamento (77,5%) é verificado quando o rótico está em posição final (P.R. .71). Do total de 861 apagamentos constatados, 626 (73%) ocorrem em fronteira de palavra (**açúcar** [a. 'su.kɐ], **fazer** [fa. 'ze], **maior** [maɪ. 'ɔ]) e apenas 232 (27%) são constatados no interior dos vocábulos (**turma** ['tu.mɐ], **perto** ['pɛ.tu], **porque** [pu. 'ke]), conforme o gráfico na Figura 3. Como apontado pelos estudos de Brandão *et al.* (2017), Brandão e De Paula (2018), Brandão (2018) e Bouchard (2017) a posição final é um contexto que favorece, consideravelmente, o apagamento do rótico em coda enquanto a posição medial o desfavorece (P.R. .29).

FIGURA 3 – Ocorrências de apagamento do rótico em coda medial e final



Fonte: elaboração própria

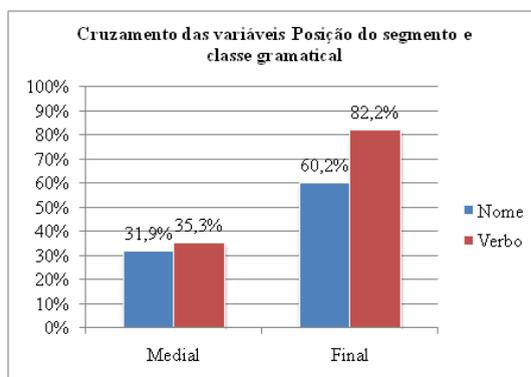
No que se refere à Escolaridade, o apagamento do rótico é favorecido por falantes com menor grau de escolaridade (P.R. .69) e desfavorecido por aqueles com ensino superior (P.R. .39) e ensino médio (P.R. .40), isto é, a probabilidade de um indivíduo mais escolarizado realizar o apagamento é bem menor do que a de um indivíduo menos escolarizado. Quanto à Faixa Etária, o maior índice de apagamento (66,1%) encontra-se entre os falantes mais jovens, com idade entre 18

¹¹ Diferença entre o menor e o maior peso relativo dos fatores de uma variável. Quanto maior o range mais significativa é a variável, uma vez que, os fatores favorecedores da aplicação da regra em análise estão mais próximos de 1 e aqueles que a desfavorecem estão mais próximos de 0.

e 20 (P.R. .68), e os menores índices estão entre os falantes de 21 a 40 anos (P.R. .42) e de 41 a 52 anos (P.R. .39), indicando que estes são mais conservadores.

A quarta variável selecionada é a Classe Gramatical da palavra que contém o rótico em coda, e o maior índice de apagamento (73,8%) ocorre nos verbos (P.R. .57), como em **comer** [ko.'me], em contraponto ao índice verificado nas palavras não-verbais (P.R. .43), tais como **turma** ['tu.mɐ]. Por fim, o sexo dos informantes é a última variável selecionada como relevante, e aponta que as mulheres favorecem o apagamento (P.R. .55), ao passo que os homens (P.R. .45) o desfavorecem. Levando em consideração que estudos sobre o mesmo fenômeno em variedades do PB têm apresentado índices quase categóricos de apagamento do rótico em coda final de itens verbais (CALLOU, 1987 apud OLIVEIRA, 2018, p. 26; CALLOU; LEITE; MORAES, 2002; CALLOU; SERRA, 2012; CALLOU; SERRA; CUNHA, 2015; OLIVEIRA, 2018; OUSHIRO; MENDES, 2014), realizamos o cruzamento das variáveis: Posição do Segmento e Classe Gramatical da palavra, conforme o gráfico na Figura 4.

FIGURA 4 – Cruzamento das variáveis Posição do Segmento e Classe Gramatical da palavra



Fonte: elaboração própria

O gráfico na Figura 4 mostra que o maior índice de apagamento é constatado em coda final de itens verbais (82,2%), como no PB. Contudo, esse índice também é bem elevado em coda final de itens não verbais (60,2%) e o que mais se destaca é a diferença entre a frequência de aplicação da regra de apagamento nas posições final e medial.

Considerando tal diferença, a amostra, de 1523 ocorrências, foi dividida em duas amostras menores: 712 ocorrências do rótico em coda medial e 811 em coda final. Feito isso, analisamos as duas amostras separadamente. Nas duas análises foi averiguada a relevância das mesmas variáveis testadas na análise inicial: Grau de Escolaridade, Faixa Etária e Sexo dos informantes, Classe Gramatical da palavra, Tonicidade da sílaba e Contexto Fonológico Precedente quanto a (i) altura [-alto] [+alto] e (ii) anterioridade [-anterior] [+anterior]. O quadro 2 resume os resultados das duas análises, evidenciando a importância das variáveis sociais para o apagamento do rótico nos dois contextos, uma vez que, em ambos, duas das três variáveis sociais testadas foram selecionadas como relevantes, enquanto nenhuma das variáveis linguísticas se mostrou relevante no contexto de coda medial, e em contexto de coda final, apenas uma das variáveis linguísticas se revelou significativa.

QUADRO 2 – Resultados da análise de apagamento do rótico em coda no PST nas posições: medial e final

Apagamento do rótico em coda medial	Apagamento do rótico em coda final
Dados analisados: 712 ocorrências	Dados analisados: 811 ocorrências
Apagamento: 32,6%	Apagamento: 77,5%
Variáveis selecionadas:	Variáveis selecionadas:
<i>Escolaridade</i>	<i>Escolaridade</i>
<i>Sexo</i>	<i>Faixa Etária</i>
	<i>Classe Gramatical</i>

Fonte: elaboração própria.

Na análise do apagamento do rótico em coda medial, o índice de apagamento foi de 32,6% (232/712) e as variáveis selecionadas como relevantes à aplicação da regra de apagamento foram: Escolaridade e Sexo dos informantes. Os resultados para os fatores das variáveis selecionadas foram semelhantes aos da análise inicial (cf. TABELAS 1 e 2) e aos de Brandão *et al.* (2017). Os maiores índices de apagamento estão entre os informantes do ensino fundamental (P.R. .62) e do sexo feminino (P.R. .59) e os menores entre os falantes com ensino superior (P.R. .27) e do sexo masculino (P.R. .41).

TABELA 2 – Variáveis atuantes no apagamento do rótico em coda medial no PST

Variável	Fatores	Apl./N	% R-Ø	P.R.
Escolaridade	Fundamental	78/175	44,6	.62
	Médio	115/290	39,6	.61
	Superior	<u>39/247</u>	15,8	.27
		232/712		<i>Range .35</i>
Sexo	Feminino	143/373	38,3	.59
	Masculino	89/339	26,2	.41
		232/712		<i>Range .18</i>
<i>Input= 0.30</i>				<i>p < 0,0001</i>

Fonte: elaboração própria.

Analisando o apagamento do rótico em coda final, contexto em que o fenômeno ocorre com muito mais frequência, sendo verificado em 77,5% (629/811) dos casos, o modelo estatístico selecionou as seguintes variáveis: Escolaridade, Faixa Etária e Classe Gramatical.

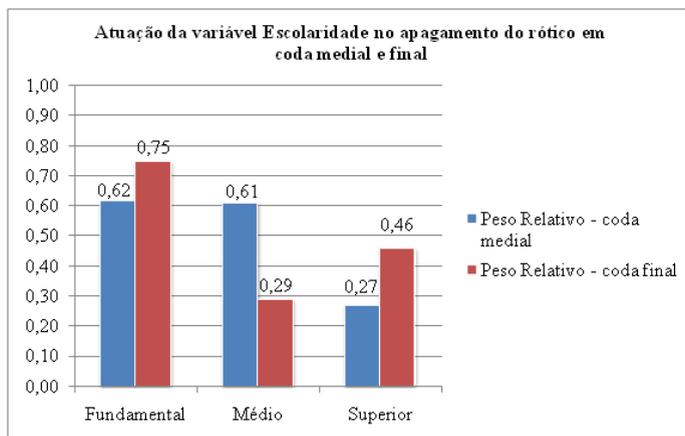
TABELA 3 – Variáveis atuantes no apagamento do rótico em coda final no PST

Variável	Fatores	Apl./N	% R-Ø	P.R.
Escolaridade	Fundamental	156/176	88,6	.75
	Médio	288/364	79,1	.29
	Superior	<u>185/271</u>	68,3	.46
		629/811		<i>Range .46</i>
Faixa Etária	18-20 anos	254/294	86,4	.76
	21-40 anos	236/329	71,7	.35
	41-52 anos	<u>139/188</u>	73,9	.37
		629/811		<i>Range .41</i>
Classe Gramatical	Verbos	526/640	82,2	.62
	Não-verbos	<u>103/171</u>	60,2	.38
		629/811		<i>Range .24</i>
<i>Input = 0.78</i>				<i>p < 0,0001</i>

Fonte: elaboração própria.

A variável Escolaridade se mostrou relevante nos dois contextos analisados de forma semelhante, sendo a mais saliente para o apagamento em posição final, com um *Range* de .46. O gráfico na Figura 5 mostra uma comparação entre os pesos relativos dos fatores dessa variável (Fundamental, Médio e Superior) na análise do apagamento do rótico em coda medial e final. Falantes com ensino fundamental são os que mais apagam o rótico tanto em coda medial (P.R. .62) quanto em coda final (P.R. .75). Todavia, o menor índice de apagamento em coda medial é o dos falantes com ensino superior (P.R. .27), enquanto os falantes do ensino médio (P.R. .61) favorecem o fenômeno quase na mesma medida que os falantes do ensino fundamental. Já em coda final, apesar de os falantes com ensino superior (P.R. .46) desfavorecerem o apagamento, o menor índice é observado entre os falantes do ensino médio (P.R. .29). Contudo, de maneira geral esses resultados indicam, semelhantemente à análise inicial, que indivíduos mais escolarizados aplicam a regra de apagamento com menos frequência do que os menos escolarizados, o que também se verifica em Brandão *et al.* (2017), Brandão e De Paula (2018), Brandão (2018) e Bouchard (2017), e pode estar associado ao fato de que falantes com maior grau de escolaridade se aproximam mais da norma de prestígio: a variedade europeia, disseminada pela escola, que possui baixos índices de apagamento do rótico (BRANDÃO; MOTA; CUNHA, 2003; MATEUS; RODRIGUES, 2003).

FIGURA 5 – Atuação da variável Escolaridade no apagamento do rótico em coda medial e final



Fonte: elaboração própria.

Na Tabela 3, verifica-se que a Faixa Etária foi a segunda variável mais saliente, com um *Range* de .41, apontando, como na análise inicial, que falantes mais jovens, entre 18 e 20 anos (P.R. .76) são mais propensos ao apagamento, ao passo que, os mais velhos, entre 21 e 40 anos (P.R. .35) e entre 41 e 52 anos (P.R. .37), são mais conservadores, não sendo verificada uma grande diferença entre os índices dessas duas últimas faixas etárias.

A Classe Gramatical foi a única variável linguística selecionada como relevante para o apagamento do rótico em coda final e, como constatado, também em Brandão *et al.* (2017), Brandão e De Paula (2018), Brandão (2018) e Bouchard (2017), o maior índice de apagamento ocorre nos verbos (P.R. .62). Em itens verbais o fenômeno ocorre em 82,2% (526/640) dos casos.

No que diz respeito aos índices percentuais de apagamento em coda medial (32,6%) e final (77,5%) é interessante notar que o índice de apagamento em coda final verificado por Bouchard (2017) foi de 79,4%, bem próximo ao que verificamos nessa mesma posição. Já nos trabalhos de Brandão *et al.* (2017), Brandão e De Paula (2018) e Brandão (2018), baseados num *corpus* de 2009, os índices de apagamento medial e final foram, respectivamente, 4,4% e 44,7%, valores bem menores do que os constatados aqui a partir da análise de um *corpus* de 2016 e 2019, e por Bouchard (2017), que analisou um *corpus* coletado entre 2015 e 2017. Isso indica que, em um intervalo de menos de dez anos, houve uma mudança considerável na proporção de implementação do fenômeno, o que, em conjunto com a relevância da variável faixa etária na análise geral e em coda final, sugere uma mudança linguística em curso (cf. WEINREICH; LABOV; HERZOG, 1968).

Considerando esses resultados em comparação com o PB e o PE, destacamos que, como o PST, diversas variedades do PB também possuem índices elevados de apagamento do rótico, enquanto o PE possui índices menores¹² (ALVES, 2015; BRANDÃO; MOTA; CUNHA, 2003; CALLOU, 1987 *apud* OLIVEIRA, 2018, p. 25-26; CALLOU;

¹² Um trabalho que evidencia essa diferença é o de Brandão, Mota e Cunha (2003) que, realizando um estudo contrastivo entre o comportamento do rótico em final de vocábulo no PE e no PB, chega a resultados opostos para as variedades brasileira e europeia no que diz respeito aos índices de apagamento: 26% de apagamento e 74% de manutenção no PE e 78% de apagamento e 22% de manutenção no PB.

LEITE; MORAES, 1994; CALLOU; SERRA, 2012; CALLOU; SERRA; CUNHA, 2015; MATEUS; RODRIGUES, 2003; MONGUILLHOT, 1997; OLIVEIRA, 2018; OUSHIRO; MENDES, 2014; SERRA; CALLOU, 2015).

Além disso, a diferença entre os índices de apagamento do rótico em coda medial e final, constatada no PST, também é verificada nas variedades no PB (ALVES, 2015; BRANDÃO; MOTA; CUNHA, 2003; GREGIS, 2001; MONARETTO, 2000; MONGUILLHOT, 1997; OUSHIRO; MENDES, 2014; PIMENTEL, 2003) e trabalhos que ultrapassam o domínio da sílaba apontam que esse seria um indício de que, ainda que o apagamento do rótico seja condicionado pela coda, sua produtividade envolveria domínios prosódicos maiores, como a fronteira de palavra prosódica (CALLOU; SERRA, 2012; CALLOU; SERRA; CUNHA, 2015; OLIVEIRA, 2018). No PE, entretanto, segundo Mateus e Rodrigues (2003), o rótico é apagado com mais frequência em posição medial, quando antecede uma consoante oclusiva ou fricativa, já quando é seguido de vogal ou pausa raramente ocorre o apagamento.

Como no PST, em variedades do PB, os índices de apagamento do rótico em coda final são mais expressivos em verbos, do que em não-verbos, sendo, inclusive, constatados índices superiores a 80% como se verificou para o PST (CALLOU, 1987 *apud* OLIVEIRA, 2018, p. 26; CALLOU; LEITE; MORAES, 2002; CALLOU; SERRA, 2012; CALLOU; SERRA; CUNHA, 2015; OLIVEIRA, 2018; OUSHIRO; MENDES, 2014), todavia, o mesmo não acontece em variedades do PE. Brandão, Mota e Cunha (2003) constataram índices quase idênticos para verbos e nomes em coda final, e Mateus e Rodrigues (2003), analisando o apagamento em coda medial e final, constataram exatamente os mesmos índices de apagamento em verbos e não-verbos.

Quanto aos resultados relativos à atuação das variáveis sociais, apresentados neste artigo, encontramos similaridades em estudos sobre variedades do PB. Os maiores índices de apagamento, geralmente, estão entre: (i) os falantes mais jovens (BRANDÃO; MOTA; CUNHA, 2003; CALLOU, 1987 *apud* OLIVEIRA, 2018, p. 25; CALLOU; SERRA; CUNHA, 2015; MONARETTO, 2000; OUSHIRO; MENDES, 2014; PIMENTEL, 2003); (ii) menos escolarizados (ALVES, 2015; BRANDÃO; MOTA; CUNHA, 2003; MONARETTO, 2000; MONGUILLHOT, 1997); e (iii) do sexo feminino (CALLOU, 1987 *apud* OLIVEIRA, 2018, p. 25; GREGIS, 2001). Já no caso do PE, Mateus e Rodrigues (2003) afirmam

que o apagamento se deve mais a fatores linguísticos do que sociais. Em Brandão, Mota e Cunha (2003), apenas a variável social Faixa Etária se mostrou relevante e, como no PB e no PST, os informantes mais jovens apresentaram o maior índice de apagamento. Para as autoras, tanto em PB quanto em PE o apagamento do (R) é uma inovação, que, no PB, já está bastante disseminada (BRANDÃO; MOTTA; CUNHA, 2003).

Em suma, o apagamento do rótico em coda, como em **parto** ['par.tu] ~ ['pa.tu] e **chamar** [ʃa.'mar] ~ [ʃa.'ma], é influenciado por variáveis sociais e linguísticas, no entanto, esse fenômeno, apesar de ser produtivo no PE, PB e PST, é muito mais recorrente nas duas últimas, sugerindo uma tendência dessas duas variedades ao padrão silábico CV, haja vista que a elisão da coda – a realização de sílabas CVC como CV – é comum em muitas outras línguas (SELKIRK, 1982), o que pode ser relacionado ao fato de que a estrutura silábica CV é universal. Essa hipótese deve, no entanto, ser avaliada em conjunto com a análise de outros *templates* silábicos do PST, observando: (i) se o apagamento em coda é recorrente, também, tendo outros segmentos licenciados neste constituinte como alvo (BALDUINO; VIEIRA, 2020; VIEIRA; BALDUINO, 2020) em itens como **susto** ['suʃ.tu] e **palma** ['pał.mə], cujas codas correspondem a /S/ e /l/, e/ou (ii) se o fenômeno é implementado na dissolução de *onsets* complexos em palavras como **prato** ['pra.tu], em que o rótico ocupa a segunda posição do *cluster*. Ademais, fatores como a debilidade do rótico, especialmente, em fronteira de palavra, um domínio da prosódia, que parece caracterizar não só o PB, mas também o PST, também devem ser considerados.

6 Considerações Finais

Na análise inicial o apagamento do rótico no PST foi constatado em 56,53% (861/1523) das ocorrências analisadas, sendo que os maiores índices de apagamento foram verificados quando o rótico estava em posição final (P.R. .71), em verbos (P.R. .57) e entre indivíduos menos escolarizados (P.R. .69) e mais jovens (P.R. .68). Na análise do fenômeno em coda medial, o índice de apagamento foi de 32,6% (232/712) com os maiores índices de apagamento sendo verificados entre os informantes do ensino fundamental (P.R. .62) e do sexo feminino (P.R. .59). Na análise do apagamento em coda final, o índice de apagamento foi de 77,5% (629/811) e os falantes do ensino fundamental (P.R. .75) e mais jovens (P.R. .76)

foram os que mais apagaram o rótico. Quanto à Classe Gramatical, o maior índice de apagamento (82,2%) ocorreu nos verbos (P.R. .62).

A diferença entre os índices de apagamento verificados no presente estudo (32,6% em coda medial e 77,5% em coda final) e os índices (4,4% em coda medial e 44,7% em coda final) constatados nos trabalhos de Brandão *et al.* (2017), Brandão e De Paula (2018) e Brandão (2018), baseados num *corpus* de 2009, indica uma possível mudança linguística em curso (cf. WEINREICH; LABOV; HERZOG, 1968). Essa hipótese é reforçada pela relevância da variável Faixa Etária na análise geral e em coda final observada neste estudo, fato também apontado por Bouchard (2017), que trabalha com um *corpus* mais recente, coletado entre 2015 e 2017.

Comparando os resultados aqui obtidos para o PST com estudos sobre variedades do PB e do PE, concluímos que as variáveis sociais e linguísticas, geralmente, se comportam de forma semelhante no PST e no PB, com índices de apagamento expressivos. Já o PE apresenta índices de apagamento menores, e as variáveis linguísticas se comportam de forma diferente nas variedades do PE analisadas. Enquanto no PB e no PST o índice de apagamento é mais expressivo em coda final e em itens verbais, no PE não há diferenças entre os índices em itens verbais e não-verbais, e o fenômeno ocorre com mais frequência em posição medial, quando precede uma consoante oclusiva ou fricativa.

Assim, concluímos que o fenômeno, apesar de ser produtivo nas três variedades, é mais recorrente no PB e no PST, sugerindo uma preferência dessas duas variedades ao padrão silábico CV, comum em muitas outras línguas (SELKIRK, 1982), e que pode ser relacionado ao fato de a estrutura silábica CV ser universal – hipótese que precisa ser avaliada em conjunto com outros processos do PST que tenham a sílaba como domínio. Ademais, fatores, como a debilidade do rótico, especialmente, em fronteira de palavra, também devem ser considerados. Dessa forma, esse estudo, não somente corrobora os resultados de trabalhos prévios como os de Bouchard (2017), Brandão *et al.* (2017), Brandão e De Paula (2018) e Brandão (2018) quanto ao comportamento das variáveis sociais e linguísticas, como também aponta a necessidade de uma análise que discuta os padrões silábicos do PST à luz de outros fenômenos fonológicos, os quais podem não somente ser afetados pelo contato com o santome e outras línguas locais, como podem sugerir inovações em relação às estruturas silábicas previstas em outras variedades da língua portuguesa.

Agradecimento

Agradecemos à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo financiamento que permitiu a condução dessa pesquisa e elaboração deste artigo: processos 2015/25332-1 e 2017/26595-1.

Contribuições de cada autora

Nancy Mendes Torres Vieira: Conceptualização; Curadoria de dados; Escrita – original; Metodologia; Análise formal; Escrita – análise e edição; Recursos.

Amanda Macedo Balduino: Conceptualização; Investigação; Coleta e curadoria de dados; Metodologia; Escrita – análise e edição; Recursos.

Referências

ABAURRE, M. B. M.; SÂNDALO, M. F. S. Os róticos revisitados. In: HORA, D.; COLLISCHONN, G. (org.). *Teoria linguística: Fonologia e outros temas*. João Pessoa: Editora da Universidade Federal da Paraíba, 2003. p. 144-180.

AGOSTINHO, A. L. *Fonologia e Método Pedagógico do Lung'ie*. 2015. 446f. Tese (Doutorado em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, 2015.

AGOSTINHO, A. L. Róticos em contexto intervocálico no Português da Ilha do Príncipe: fonologia e educação. In: ENCONTRO DA ABECS, IX., 2016, Brasília. *Anais* [...]. Brasília: Universidade de Brasília, 2016. p. 48-49. Disponível em: <https://encontroabecs.wordpress.com/cad-de-resumos>. Acesso em: 10 set. 2020

ALVES, M. A. Variação na produção/apagamento da vibrante pós-vocálica no falar florianopolitano. *Working Papers em Linguística*, Florianópolis, v. 16, n. 1, p. 20-35, 2015. DOI: <https://doi.org/10.5007/1984-8420.2015v16n1p20>.

ARAÚJO, G. Há uma política linguística para o português em São Tomé e Príncipe?. In: SOUZA, S.; OLMO, F. C. (org.). *Línguas em português: a lusofonia numa visão crítica*. Porto: Universidade do Porto Press, 2020. p. 173-197.

ARAÚJO, G.; HAGEMEIJER, T. *Dicionário Santome-Português/Português-Santome*. São Paulo: Hedra, 2013.

BALDUINO, A. M. *A nasalidade vocálica no português falado em São Tomé e Príncipe*. 2018. 296f. Dissertação (Mestrado em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, 2018.

BALDUINO, A. M. Apagamento de /R/ e /S/ em coda no Português Principense. *Papia*, São Paulo, v. 29, n. 1, p. 25-39, 2019. Disponível em: http://revistas.ffch.usp.br/papia/article/view/3366/pdf_1. Acesso em: 23 jul. 2020.

BALDUINO, A. M. *Processos fonológicos no português de São Tomé e de Santo António do Príncipe*. Tese (Doutorado em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo. (em preparação).

BALDUINO, A. M.; VIEIRA, N. M. T. Distribuição da lateral /l/ em coda no português santomense. *Estudos Linguísticos*, São Carlos, SP, v. 49, n. 2, p. 594-615, 2020. DOI: <https://doi.org/10.21165/el.v49i2.2490>.

BANDEIRA, M. *Reconstrução fonológica e lexical do protocrioulo do Golfo da Guiné*. 2017. 439f. Tese (Doutorado em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2017.

BAXTER, A. O Português dos Tongas de São Tomé. In: OLIVEIRA, M. D.; ARAUJO, G. (org.). *O português na África Atlântica*. São Paulo: Humanitas/Fapesp, 2018. p. 297-324.

BISOL, L. A sílaba e seus constituintes. In: NEVES, M. H. M (org.). *Gramática do Português culto falado: novos estudos*. Campinas: Editora da Unicamp, 1999. p. 701-742.

BOERSMA, P.; WEENINK, D. *Praat: doing phonetics by computer* (Version 5.3.82) Computer Program. 2020. Disponível em: <http://www.praat.org>. Acesso em: 20 jun. 2020.

BOUCHARD, M. E. *Linguistic Variation and Change in the Portuguese of São Tomé*. 2017. 389f. Dissertation (Doctoral) – Department of Linguistics, New York University, New York, 2017. DOI: <https://doi.org/10.5334/jpl.192>

BRANDÃO, S. F. Apagamento de R em coda externa em duas variedades africanas do português. *Diadorim*, Rio de Janeiro, v. 20, n. especial, p. 390-408, 2018. DOI: <https://doi.org/10.35520/diadorim.2018.v20n0a23283>.

BRANDÃO, S. F.; MOTA, M. A.; CUNHA, C. S. Um estudo contrastivo entre o português europeu e o português do Brasil: o –R final de vocábulo. In: BRANDÃO, S. F.; MOTA, M. A. (org.). *Análise contrastiva de variedades do português: primeiros estudos*. Rio de Janeiro: In-fólio, 2003. p. 163-180.

BRANDÃO, S. F.; DE PAULA, A. Róticos nas variedades santomense e moçambicana do Português. In: BRANDÃO, S. F. (org.). *Dois variedades africanas do Português: variáveis fonético-fonológicas e morfossintáticas*. São Paulo: Blucher, 2018. p. 93-118. DOI: <https://doi.org/10.5151/9788580393248>

BRANDÃO, S. F.; PESSANHA, D. B.; PONTES, S.; CORRÊA, M. Róticos na variedade urbana do Português de São Tomé. *Papia*, São Paulo, v. 27, n. 2, p. 293-315, 2017. Disponível em: <http://revistas.fflch.usp.br/papia/article/view/2762/pdf>. Acesso em: 20 jul. 2020.

CALLOU, D.; LEITE, Y.; MORAES, J. Variação e diferenciação dialetal: a pronúncia do /r/ no português do Brasil. In: KOCH, I. G. V. (org.). *Gramática do português falado*. Campinas: Editora da Unicamp/FAPESP, 1994. p. 465-493.

CALLOU, D.; LEITE, Y.; MORAES, J. Processo(s) de enfraquecimento consonantal no português do Brasil. In: ABAURRE, M. B. M.; RODRIGUES, A. C. S. (org.). *Gramática do português falado VIII: novos estudos descritivos*. Campinas: UNICAMP/FAPESP, 2002. p. 537-555.

CALLOU, D.; SERRA, C. Variação do rótico e estrutura prosódica. *GELNE*, Natal, v. 14, n. especial, p. 41-58, 2012. Disponível em: <https://periodicos.ufrn.br/gelne/article/view/9363>. Acesso em: 20 jul. 2020.

CALLOU, D.; SERRA, C.; CUNHA, C. Mudança em curso no português brasileiro: o apagamento do R no dialeto nordestino. *Revista Abralín*, Curitiba, v. 14 n. 1, p. 195-219, 2015. DOI: <http://dx.doi.org/10.5380/rabl.v14i1.42491>.

CÂMARA Jr., J. M. *Estrutura da língua portuguesa*. Petrópolis: Vozes, 1970.

CHESHIRE, J. Sex and Gender in Variationist Research. In: TRUDGILL, P.; CHAMBERS, J. K.; SCHILLING-ESTES, N. (org.). *The Handbook of Language Variation and Change*. Oxford: Blackwell Publishing, 2004. p. 423-443. DOI: <https://doi.org/10.1111/b.9781405116923.2003.00024.x>

CHRISTOFOLETTI, A. *Ditongos no português de São Tomé e Príncipe*. 2013. 109f. Dissertação (Mestrado em Letras) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2013.

FERRAZ, L. *The Creole of São Tomé*. Johannesburg: Witwatersrand University Press, 1979.

FIGUEIREDO, C. *A concordância plural variável no sintagma nominal do português reestruturado da comunidade de almoxarife, São Tomé*. 2010. 792f. Tese (Doutorado em Linguística) – Faculdade de Ciências Sociais e Humanas, Universidade de Macau, Macau, 2010.

GOMES, C. A. Para além dos pacotes estatísticos Varbrul/Goldvarb e Rbrul: qual a concepção de gramática? *GELNE*, Natal, v. 14, n. especial, p. 259-272, 2012. Disponível em: <https://periodicos.ufrn.br/gelne/article/view/9363>. Acesso em: 20 jul. 2020.

GONÇALVES, R. *Propriedade de subcategorização verbal no português de S. Tomé*. 2010. 151f. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade de Lisboa, Lisboa, 2010.

GONÇALVES, R.; HAGEMEIJER, T. O português num contexto multilingue: o caso de São Tomé e Príncipe. *Revista Científica da Universidade Eduardo Mondlane*, Maputo, v. 1, n. 1, p. 84-103, 2015. Disponível em: <https://repositorio.ul.pt/bitstream/10451/31032/1/Goncalves%26Hagemeijer2015.pdf>. Acesso em: 25 jul. 2020.

GREGIS, H. *O apagamento da vibrante pós-vocálica em Porto Alegre*. 2001. 116f. Dissertação (Mestrado em Letras) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2001.

GUY, G.; ZILLES, A. *Sociolinguística quantitativa*. São Paulo: Parábola Editorial, 2007.

HAGEMEIJER, T. From Creoles to Portuguese: Language Shift in São Tomé and Príncipe. In: LÓPEZ, L.; GONÇALVES, P.; AVELAR, J. (org.). *The Portuguese Language Continuum in Africa and Brazil*. Amsterdam; Philadelphia: John Benjamins, 2018. p. 169-184. DOI: <https://doi.org/10.1075/ihll.20.08hag>

INE - Instituto Nacional de Estatística de São Tomé e Príncipe. 2012. Disponível em: <https://www.ine.st/>. Acesso em: 10 jun. 2020.

LABOV, W. *Sociolinguistic Patterns*, Philadelphia: University of Pennsylvania Press, 1972.

LABOV, W. *Principles of Linguistic Change: Internal Factors*. Oxford: Blackwell, 1994. v. 1.

LABOV, W. *Principles of Linguistic Change: Social Factors*. Oxford: Blackwell, 2001. v. 2.

MATEUS, M. H.; D'ANDRADE, E. *The Phonology of Portuguese*. Oxford: Oxford Linguistics, 2000.

MATEUS, M. H. M.; RODRIGUES, C. A Vibrante em Coda em Português Europeu. In: HORA, D.; COLLISCHONN, G. (org.). *Teoria Linguística Fonologia e outros temas*. João Pessoa: Ed. Universitária, 2003. p.181-199.

MENESES, F. O. *As vogais desvozeadas no Português Brasileiro: investigação acústico-articulatória*. 2012. 123f. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2012.

MONARETTO, V. N. O. O apagamento da vibrante pós-vocálica nas capitais do sul do Brasil. *Letras de Hoje*. Porto Alegre, v. 35, n. 1, p. 275-284, 2000. Disponível em: <https://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/14768/9834>. Acesso em: 20 jul. 2020.

MONGUILHOT, I. O. S. A vibrante em final de palavra na fala de Santa Catarina. In: ENCONTRO DO CÍRCULO DE ESTUDOS LINGÜÍSTICOS DO SUL, 2., 1997, Florianópolis. *Anais do II CELSUL*. Florianópolis: Universidade Federal de Santa Catarina, 1997. [1 CD-Rom]

OLIVEIRA, I. C. *Os róticos em coda silábica externa: o interior da região Sul no projeto AliB*. 2018. 133f. Dissertação (Mestrado em Letras) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2018.

OUSHIRO, L. Tratamento de dados com o R para análises sociolinguísticas. In: FREITAG, R. M. K. (org.). *Metodologia de coleta e manipulação de dados em Sociolinguística*. São Paulo: Editora Blucher, 2014. p. 133-177. DOI: <https://doi.org/10.5151/BlucherOA-MCMDs-10cap>

OUSHIRO, L. *Introdução à Estatística para Linguistas*, v.1.0.1. 2017. Disponível em: <https://rpubs.com/oushiro/iel>. Acesso em: 2 jul. 2020.

OUSHIRO, L.; MENDES, R. B. O apagamento de (-r) em coda nos limites da variação. *Veredas*, Juiz de Fora, v. 18, n. 2, p. 251-266, 2014. Disponível em: <https://periodicos.ufjf.br/index.php/veredas/article/view/24963>. Acesso em: 2 jul. 2020.

PIMENTEL, R. M. *A variação linguística do fonema /r/ na posição pós-vocálica*. 2003. 105f. Dissertação (Mestrado em Letras) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2003.

QUEDNAU, L. *A lateral pós-vocálica no português gaúcho: análise variacionista e representação não linear*. 1993. 110f. Dissertação (Mestrado em Letras) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1993.

RODRIGUES, M. C. Todas as codas são frágeis em português europeu? *Revista Linguística*, Rio de Janeiro, v. 8, n. 1, p. 138-149, 2012.

ROUGÉ, J. L. Les langues des Tonga. In: D'ANDRADE, E.; KIHM, A. (org.). *Actas do Colóquio sobre Crioulos de base lexical portuguesa*. Lisboa: Colibri, 1992. p. 171-76.

SANTIAGO, A. M. *As vogais do português do Príncipe*. Trabalho de conclusão de curso (Graduação em Letras) – Centro de Comunicação e Expressão, Universidade Federal de Santa Catarina, Florianópolis, 2019.

SANTIAGO, A. M.; AGOSTINHO, A. L. Situação linguística do português em São Tomé e Príncipe. *A cor das Letras*, Feira de Santana, BA, v. 21, n. 1, p. 39-61, 2020. DOI: <https://doi.org/10.13102/cl.v21i1.4970>.

SELKIRK, E. The Syllable. In: HULST, H.; SMITH, N. (ed.). *The Structure of Phonological Representations*. Dordrecht: Foris, 1982. p. 337-383.

SERRA, C.; CALLOU, D. Prosodic Structure, Prominence and /r/-Deletion in Final Coda Position: Brazilian Portuguese and European Portuguese Contrasted. In: DOMINICIS, A. (org.). *pS-prominenceS: Prominences in Linguistics*. Viterbo: Disucom Press, 2015. p. 96-113.

VIEIRA, N. M. T.; BALDUINO, A. M. Apagamento de /R, S, l/ na coda no português de São Tomé: convergência linguística? *Papia*, São Paulo, v. 30, n. 1, p. 7-33, 2020. Disponível em: <http://www.revistas.fffch.usp.br/papia/article/view/3415/pdf>. Acesso em: 3 set. 2020.

WEINREICH, U.; LABOV, W.; HERZOG, M. Empirical Foundations for a Theory of Linguistic Change. In: LEHMANN, W. P.; MALKIEL, Y. (org.). *Directions for Historical Linguistics*. Austin: University of Texas Press, 1968. p. 97-195.