

A CONSTRUÇÃO E A RELEVÂNCIA DE UM *CORPUS* COM DADOS DO PORTUGUÊS ESCRITO CULTO RORAIMENSE¹

Lana Camila Santos Gonçalves²
Caroline Carnielli Biazolli³

Resumo: Este estudo busca debater a importância da construção de *corpora* para o avanço das investigações (sócio)linguísticas de Roraima – que possui as suas pesquisas em pleno desenvolvimento. Para tanto, baseia-se no *corpus* de Gonçalves (2024) – em andamento – para discutir o cenário local e os critérios metodológicos utilizados para a construção das amostras. Esse material oportuniza tamanha discussão porque evidencia a organização de gêneros textuais-discursivos de um jornal reconhecido como o de maior credibilidade do estado – a Folha de Boa Vista. Este trabalho resulta no primeiro *corpus* com foco no português escrito culto de Roraima e na ampliação do banco de dados linguísticos local, o que encaminha o estado rumo a uma formação completa de materiais que tratem da língua. O material construído será disponibilizado em breve à comunidade acadêmica em um repositório preparado para recebê-lo.

Palavras-chave: Gêneros textuais-discursivos jornalísticos; *Corpus*; Português escrito culto; Português de Roraima; Sociolinguística.

The construction and relevance of a *corpus* with data from cultured written Portuguese from Roraima

Abstract: The present study aims to debate the importance of constructing *corpora* to advance the (socio)linguistic investigations in Roraima – where research is in full development. In order to do so, the *corpus* by Gonçalves

¹ Este trabalho conta com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

² Universidade Federal de São Carlos. Orcid: <https://orcid.org/0000-0002-2942-3037>, E-mail: lanagoncalves@estudante.ufscar.br.

³ Universidade Federal de São Carlos. Orcid: <https://orcid.org/0000-0002-8578-8102>, E-mail: caroline.biazolli@ufscar.br.

(2024) – ongoing – is used to discuss the local scenario and the methodological criteria employed to construct the samples. This material allows such discussion because it shows the organization of genres from a newspaper recognized as having the most credibility in the state – *Folha de Boa Vista*. This research results in the first *corpus* focused on the cultured written Portuguese from Roraima and the broadening of local linguistic databases, which sets the state on a way towards the creation of complete materials concerning language. The constructed material will soon be made available to the academic community in a repository prepared to receive it. **Keywords:** Journalistic genres; *Corpus*; Cultured written Portuguese; Portuguese from Roraima; Sociolinguistics.

INTRODUÇÃO

Este estudo apresenta um recorte de uma dissertação (Gonçalves, 2024), ainda em desenvolvimento, que, até o presente momento, organizou amostras de textos jornalísticos na intenção de intensificar a inclusão do português de Roraima (PRR) nas discussões nacionais sobre a realização do português brasileiro, visando, posteriormente, à caracterização do português escrito culto roraimense. Diante disso, este artigo objetiva debater a relevância do *corpus* constituído, e de outros *corpora* voltados às pesquisas (sócio)linguísticas locais, e compartilhar a metodologia empregada para a construção do material em foco.

A motivação para a construção de um *corpus* do PRR advém da escassez de *corpora* de dados linguísticos do estado, justificada até mesmo pela recente consolidação do território roraimense. Pesquisadores têm se movimentado cada vez mais em função da descrição e da análise de características que permeiam a fala e a escrita local. Nessa direção, já podem ser vistas propostas de *corpora* acerca da fala roraimense, tais como o desenvolvimento do “Atlas Linguístico de Roraima”, com base nas diretrizes e nos objetivos do ALiB (Atlas Linguístico Brasileiro), sob a responsabilidade do NEPSol-RR (Núcleo de Estudos e Pesquisas Sociolinguísticas de Roraima), coordenado pela Profa. Dra. Luzineth Rodrigues Martins, e do projeto “Constituição de amostras de fala: contribuindo para os estudos de fenômenos linguísticos no estado de Roraima”, sob a liderança da Profa. Dra. Eliaine de Moraes Belford Gomes – entre outras propostas limitadas e circunscritas a dissertações. Contudo, um trabalho voltado ao PRR escrito culto com foco em dados da língua extraídos de gêneros textuais-

discursivos⁴ jornalísticos, como estabelecido nesta pesquisa, ainda não tinha sido realizado – sendo o material aqui discutido a representação de um novo momento para os estudos linguísticos locais.

Apoiado em pressupostos da linguística de *corpus* (Sardinha, 2004) e da sociolinguística (Labov, 2008 [1972]), sobretudo no que concerne às correlações entre variação e gêneros (Biazolli, 2016; Vieira; Lima, 2019; Biazolli; Berlinck, 2021b) e à pluralidade de normas linguísticas (Faraco, 2008; Bagno, 2011), o presente trabalho pauta-se, inicialmente, em diretrizes metodológicas já estabelecidas no projeto “Pró-norma plural: do *continuum* fala-escrita para a norma-padrão” (Vieira, em andamento)⁵ e, com o avançar das etapas de reconhecer, selecionar, recortar e organizar os gêneros textuais-discursivos do jornal de Roraima selecionado como objeto da investigação, em diretrizes elaboradas de acordo com as especificidades dos textos roraimenses. A Folha de Boa Vista (FBV) é o periódico usado para a construção do *corpus*, uma vez que é identificado pelos roraimenses como o jornal de maior credibilidade (Pimentel, 1996), sendo o único que já teve circulação impressa. O grupo FBV, hoje nas plataformas digitais, faz coberturas diárias sobre todo o estado.

Este artigo está dividido, para além das palavras introdutórias e das considerações finais, em três seções. Na primeira seção, apresenta-se uma discussão sobre os (ainda poucos) estudos (sócio)linguísticos já realizados sobre o PRR, a fim de demonstrar o quanto bancos de dados linguísticos sobre essa

⁴ Este estudo adota o termo “gêneros textuais-discursivos”, uma vez que “(...) tanto a compreensão da estrutura/organização de um texto que materializa determinado gênero quanto a compreensão dos significados do(s) discurso(s) por trás desse mesmo texto têm o mesmo peso de significância” (Biazolli; Berlinck, 2021a, p. 7).

⁵ Esse projeto, que é coordenado pela Profa. Dra. Silvia Rodrigues Vieira e é desenvolvido por uma equipe nacional, subdividida em equipes regionais, “(...) busca realizar o mapeamento de fenômenos variáveis do português brasileiro em gêneros do *continuum* fala-escrita, com vistas à proposição de uma norma de referência (norma-padrão) atualizada, de caráter plural (...). Tendo em vista essa finalidade, o projeto propõe a construção de um banco de dados, selecionando textos da modalidade escrita culta brasileira provenientes de gêneros diferentes – no caso, gêneros pertencentes às instâncias discursivas jornalística e acadêmica e produzidos nas capitais Rio de Janeiro, São Paulo, Salvador e Recife –, e o estudo de regras variáveis morfossintáticas a partir desse banco de dados construído” (Biazolli *et al.*, 2023, p. 62).

variedade poderiam contribuir para o avanço de pesquisas locais. Na sequência, reforça-se a pertinência de trabalhos voltados à constituição de *corpora* com dados linguísticos roraimenses. Na terceira seção, apresenta-se o *corpus* elaborado com base em cinco gêneros textuais-discursivos presentes na FBV – o artigo de opinião, a crônica, a notícia, a reportagem e o comentário opinativo –, abordando a representatividade sociocultural desse material, bem como a metodologia empregada para se chegar a esse banco com aproximadamente 101.300 palavras.

O CENÁRIO ATUAL DE RORAIMA: IMPULSIONAMENTO PARA OS ESTUDOS (SÓCIO)LINGUÍSTICOS

Em comparação a outros estados, os estudos (sócio)linguísticos ganharam mais espaço (e força) no estado de Roraima apenas recentemente⁶. A entrada de imigrantes venezuelanos – mais intensa nos anos de 2015 a 2017, pela forma emergente e simultânea que ocorreu – pode ser considerada um dos fatos históricos que mais impulsionou os estudiosos (e até mesmo a sociedade em geral) a refletirem sobre a língua empregada no estado. Quanto a isso, as atuais inserções de pesquisadores roraimenses em projetos nacionais demonstram a sua preocupação no que se refere ao avanço desses estudos na localidade, como se verá no decorrer desta seção.

Um dos estudos que foca a compreensão do falar roraimense é o de Gonçalves (2021). A autora faz uma catalogação sociodialetal de uma das cidades do interior do sul do estado, o que a faz constatar a conservação de dialetos nordestinos na região sul de Roraima. Sob uma perspectiva geográfica e histórica de Roraima, sem preocupações notadamente linguísticas, Freitas (1997) também identifica essa tendência em algumas partes da cidade de Boa Vista, ao dizer que, em certos lugares, “(...) há a predominância de um estado, por exemplo, na Feira do Produtor. Ali predomina o sotaque maranhense. No centro da cidade, em

⁶ Por exemplo, apenas no ano de 2018, foi que a disciplina Sociolinguística ingressou no currículo do principal e maior Curso de Letras do estado, o da Universidade Federal de Roraima. Uma iniciativa do Prof. Dr. Eliabe dos Santos Procópio.

especial na avenida Jaime Brasil, o predomínio é cearense” (Freitas, 1997, p. 164).

Dando continuidade ao objetivo unificado de compreender as marcas roraimenses e, em alguns casos, o conhecimento do indivíduo roraimense acerca disso, outros trabalhos já concluídos também devem ser mencionados, tais como o estudo sobre os neologismos de Roraima (Procópio; Silva, 2021; Procópio; Silva, 2022; Cruz; Procópio, 2023), a elaboração de um “Dicionário de Palavras e Expressões do Português de Roraima” (Procópio, 2021), a investigação sobre os cruzamentos vocabulares com o topônimo Roraima (Procópio; Silva, 2023) e as pesquisas sobre percepções sociolinguísticas do PRP (Nascimento, 2022; Sene; Chaves; Procópio, no prelo).

Em termos de projetos de pesquisa de caráter nacional, estão em construção o “Atlas Linguístico de Roraima”, iniciado pelo grupo NEPSol-RR – como já dito – em 2021, e a participação de Roraima no PHPB (Projeto Para a História do Português Brasileiro), sob a coordenação do Prof. Dr. Eliabe dos Santos Procópio.

Embora seja notável a atual efervescência de estudos (sócio)linguísticos no estado, um avanço ainda maior fica impossibilitado pelo fato de o contexto ainda não apresentar materiais suficientes que sirvam de fonte de extração de dados aos pesquisadores que desejam identificar, descrever e analisar o uso real da fala e da escrita dos roraimenses. Nessa conjuntura, a existência de *corpora* do PRR, que pudessem oferecer evidências (Hunston, 2002) dessa variedade, contribuiria para o reconhecimento – mais acelerado – de informações sobre o português roraimense.

A falta desses materiais para o desenvolvimento de estudos (sócio)linguísticos sobre o estado pode ser entendida a partir de duas principais justificativas: a primeira refere-se à consolidação recente do estado (Roraima possui apenas 35 anos de emancipação) e, conseqüentemente, à movimentação (sócio)linguística local também ser principiante; e, a segunda, à morosidade e ao notável dispêndio de tempo que envolvem a construção de um *corpus*, trabalho que requer critérios e rigor metodológicos, inclusive o envolvimento de uma equipe.

Para exemplificar, na constituição de *corpora* com dados da fala roraimense, a dificuldade é a localização de falantes que sejam nativos de Roraima – por ser um estado novo e ser, majoritariamente, formado por migrantes (Souza, 2006). Isso acaba direcionando o pesquisador a conhecer outros termos e a adotá-los em seus estudos – entre eles, está o termo *roraimado*, pensado pelo poeta Eliakin Rufino para designar aquele que não nasceu, mas mora em Roraima. Em certas circunstâncias, o pesquisador passa a considerar o roraimado um sujeito de duplo pertencimento apto a ser um participante de sua pesquisa – por exemplo, a ser seu entrevistado –, levando em consideração o seu estado de origem e o seu estado de acolhimento (Luz, 2013).

Vale (2007) descreve um cenário roraimense que, de fato, veio a se concretizar. Para a autora,

Roraima, com o passar dos anos, terá uma das populações mais mescladas do Brasil. Índios de várias tribos, garimpeiros de todas as regiões do Brasil, militares, pois esta é área de fronteira, colonos migrantes do País inteiro. Segundo Barros (1998, p. 135), um novo tipo de povo, não mais garimpeiro, porque as mineradoras tomarão o seu lugar. Não mais índio, pois sua cultura foi destruída. Não mais migrante colono porque a última fronteira agrícola do Brasil está se esgotando.

As características identitárias de todos estão sendo assimiladas, hibridizadas, compondo uma cultura que contribuirá na formação social do sujeito roraimense, que sem perceber acumula identidades e reformula a sua, numa complexa teia de redes sociais que se forma na pluralidade de culturas (Vale, 2007, p. 189-190).

Por outro lado, quando se trata de *corpora* com dados da modalidade escrita – neste caso, voltando-se à escrita jornalística –, a dificuldade é a criação e/ou o atendimento a todos os critérios que organizem o material como representativo da comunidade local, pelo fato de a imprensa roraimense, em determinados aspectos, ainda estar se estruturando, o que impossibilita o pesquisador, por exemplo, de aplicar aos jornais roraimenses a mesma

compreensão que se tem sobre grandes jornais já consolidados de outros estados, como a *Folha de S.Paulo*, de São Paulo, e *O Globo*, do Rio de Janeiro, entre outros.

O estudo que norteia este artigo encaixa-se neste último cenário, ao propor a construção de um banco de dados a partir de uma diversidade de textos jornalísticos da modalidade escrita culta roraimense. No *corpus* em questão, a organização de textos, além de servir para entender a composição do estado a ser analisado, também deve direcionar a reflexões sobre o comportamento jornalístico da FBV e os diferentes usos linguísticos que poderão vir a ser encontrados nas análises do material. O propósito maior é que esse material sirva para fins (sócio)linguísticos, mas não só, já que pesquisadores de outras áreas poderão se interessar pelos conteúdos trazidos em cada um dos textos, referentes aos acontecimentos dos últimos anos de Roraima.

Ciente, então, da urgência de um material (sócio)linguístico do PRR, este estudo soma-se a outros que têm apoiado e colaborado com a criação de *corpora* no estado, em função do entendimento dessa variedade e da inclusão dos estudos roraimenses nas análises do português brasileiro.

A RELEVÂNCIA DE *CORPORA* COM DADOS RORAIMENSES

Segundo Sanches (1995, p. 8-9 *apud* Sardinha, 2000, p. 338) *corpus* é

Um conjunto de dados linguísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso linguístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise.

Para exemplificar o trabalho de organizar dados reais de língua no Brasil, um dos grandes projetos a ser citado é o NURC (Norma Urbana Culta), que

completou 50 anos em 2019 e que, segundo Freitag (2019), tem a possibilidade de ser considerado um banco de dados sociolinguístico. Isso porque a autora conceitua o termo de “documentação sociolinguística” como um conjunto de procedimentos para registro e armazenamento padronizados de uma língua, seja ela falada, escrita ou sinalizada.

Na defesa desses materiais para os estudos da língua, Bagno (2019) pontua que tanto os dados do NURC quanto os de outros projetos que envolvam usos reais e autênticos do português brasileiro possibilitam aos pesquisadores chegarem, com segurança, a um conjunto robusto de afirmações sobre essa língua e suas características gramaticais, viabilizando, inclusive, a sua diferenciação do português europeu, assim como de demais línguas do grupo românico.

Em função disso, (sócio)linguistas de Roraima já têm começado a movimentar/efetivar a ideia de constituir *corpora*, sendo o primeiro, já citado anteriormente, o “Atlas Linguístico de Roraima”, pelo fato de Roraima ser um dos poucos estados do Brasil que ainda não tem esse material organizado, e o segundo, também já sinalizado, o projeto “Constituição de amostras de fala: contribuindo para os estudos de fenômenos linguísticos no estado de Roraima”. Este último, em comparação ao Atlas, que faz uso de entrevista semiestruturada, busca criar um banco de dados com falas espontâneas do povo roraimense, seguindo o modelo originário de entrevistas sociolinguísticas e possibilitando análises em diversos níveis.

Outro *corpus* que deve ser comentado é o vinculado ao projeto “Constituição de um *corpus* para o português de Roraima (SÉC. XVIII)”, proposto por Procópio (2023), com o objetivo de buscar materiais que representem a implantação da língua portuguesa em Roraima. O autor, que aponta a escassez desses materiais e defende o avanço da Filologia na Amazônia, compartilha a justificativa e os objetivos específicos do trabalho:

O estudo linguístico do *Cópus* tem se concentrado no nível lexical e toponímico. Como não havia uma população lusoparlante local, não é possível estabelecer a hipótese de constituição de um português também local e com marcas gramaticais próprias, inclusive porque os indígenas não tinham

acesso à escrita para expressar seu português de contato (Procópio, 2023, p. 8).

O *corpus* debatido neste trabalho – o das amostras de textos jornalísticos roraimenses –, assim como Procópio (2023), interessa-se também pela escrita, contudo, com o objetivo de buscar um material jornalístico mais recente para (futuras) descrições do PRR em diferentes níveis. A ideia de um *corpus* com textos escritos, além de ter uma característica temporal inédita no estado, servirá para compor o conjunto de bancos de dados linguísticos local.

Os pesquisadores locais têm visualizado a potencialidade de se trabalhar com *corpora*, visto que são inúmeras as possibilidades que eles permitirão às pesquisas do estado, sendo um ponto de partida marcante para o incentivo aos próximos interessados nas características (sócio)linguísticas roraimenses.

A APRESENTAÇÃO DO *CORPUS* COM DADOS DO PRR ESCRITO CULTO

Para mencionar a representatividade de um *corpus*, é preciso entender a especificidade da sua composição, bem como o que está sendo proposto para ser trabalhado. Entretanto, algumas questões são comuns aos *corpora*, independentemente de serem materiais formados por dados de fala ou de escrita. Uma delas, por exemplo, é o fato de que a representatividade de um *corpus* pode ser vista com base na sua extensão, incluindo diferentes manifestações da língua(gem).

Essa discussão é realizada por Stefanowitsch (2020), que assegura que a representatividade precisa ser feita de maneira que se possa refletir o quantitativo e o qualitativo das variedades da comunidade investigada. Simultaneamente, o autor também denomina o *corpus* representativo como um *corpus* equilibrado. E é justamente nesse equilíbrio que mora a problematização. Afinal, o maior desafio para um criador de *corpus* é elaborar critérios que apresentem o êxito em retratar, o máximo possível, os falantes/usuários daquela variedade investigada (a questão mais discutível na criação de *corpora* em Roraima hoje é, pontualmente, elaborar esses critérios). Ao conseguir esse feito, consegue-se, então, a conquista de um *corpus* equilibrado.

A FBV, jornal utilizado nesta pesquisa para a construção do *corpus* com dados do PRR escrito culto, foi considerada representativa do estado de Roraima por atender aos seguintes critérios: (i) **tempo de permanência**, pela sua continuidade jornalística de 40 anos de funcionamento local; (ii) **abrangência**, pela sua circulação em todos os 15 municípios do estado e pela sua presença em, praticamente, todas as plataformas digitais; (iii) **abordagens locais**, por ser possível observar, em sua política editorial, a predominância de pautas roraimenses; (iv) **origem dos escreventes**, pela participação de alguns roraimenses na construção da FBV, como a presença da jornalista Shirley Rodrigues; e (v) **instituição local**, por ser originalmente uma instituição nascida no estado.

A adoção deste último critério reconhece que parte dos jornais que já circularam no estado vieram do estado do Amazonas - como apresentado nas discussões de Correia (2021) - ou são filiais nacionais, como o G1 (portal de notícias da Globo).

A FBV, uma vez, já foi objeto de interesse de pesquisadores do NUDUCHIS (Núcleo de Documentação Histórica) da Universidade Federal de Roraima, que organizaram um acervo de textos das décadas de 1980-1990. Assim como esses pesquisadores/historiadores visualizaram a potencialidade da FBV para discussões históricas, o trabalho aqui apresentado segue essa referência, sendo o primeiro a organizar uma coleção de textos da FBV com foco (sócio)linguístico.

Para que se compreenda melhor todo esse uso social da FBV, é preciso recordar que o jornalismo de Roraima é marcado por descontinuidades, uma vez que se pode contabilizar - em média - o surgimento de quase 50 noticiários no decorrer desses mais de 100 anos de território roraimense (Correia, 2021). Estudos apontam que, durante a consolidação do estado, vários jornais foram abertos e fechados, sendo a maioria alvo de atritos políticos que resultavam até em assassinatos, detúncias e perseguições contra os donos.

Contrariando essa realidade local, nasce a FBV em 1983 como a primeira tentativa em Roraima de um jornal estruturado, que buscou se aproximar do estilo dos grandes jornais, com uma linha de produção que começa na captação da notícia e termina na distribuição de exemplares (Silva; Vieira, 2010).

Diante de sua insistência temporal, hoje a FBV é considerada um grupo familiar (GRUPO FOLHABV), pois possui vários repartimentos e a liderança de cada um deles está nas mãos de alguém da família, sendo a família do economista e ex-governador Getúlio Cruz a responsável por tudo – ele, por exemplo, é o responsável pela rádio do jornal – a Folha FM –, enquanto a filha dele, Paula Cruz, assumiu o comando da direção geral do jornal. A definição dessas lideranças (porque a FBV já foi liderada por seus outros fundadores e contribuintes) foi fazendo com que, aos poucos, o jornal estabelecesse as suas marcas no mercado.

A FBV é a principal responsável por dar dinamicidade ao jornalismo roraimense (Gonçalves, 2024), pois foi o único jornal do estado que manteve, por muito tempo, a tradição da versão impressa e foi o primeiro a criar um *website* jornalístico local (cf. as figuras 1 e 2). Nesse viés de avanços, de 2006 para cá, muitos atributos da Folha mudaram com a integração e o aperfeiçoamento do jornal em várias plataformas de comunicação. Segundo Silva e Vieira (2010), também nesse mesmo ano já se contabilizava que o alcance do jornal estava em cerca de 90% dos leitores de todo o estado, pois já tinha chegado aos 15 municípios e possuía uma média de 12 mil tiragens diárias.

Figuras 1 e 2: FolhaWeb e jornal impresso



Fonte: Site da FBV (2023)

Em 2022, a empresa segue com a sua reputação e credibilidade ao receber a Comenda “Orgulho de Roraima”, como uma forma de agradecimento pelos longos anos de serviços prestados no estado. Ou seja, há uma valorização regional do Grupo, advinda também da população do interior e da capital – por ser,

indiscutivelmente, o veículo de comunicação que circulou/circula por todas as regiões interioranas do estado.

Assim sendo, o subtítulo “um jornal necessário”, que foi criado lá atrás quando o Grupo passava por mais uma crise e cogitava a possibilidade de fechamento, sendo preciso um dos seus integrantes se posicionar, afirmando que a Folha não poderia fechar por ser “um jornal necessário”, continua a fazer sentido. Isso porque, independentemente de análises ideológicas, é o único jornal que venceu os obstáculos temporal e político, e que por esse motivo possui material para ser discutido em seu caráter social, histórico e linguístico.

A metodologia empregada para a construção do *corpus* de gêneros textuais-discursivos jornalísticos

A metodologia de um trabalho que envolve a criação de um *corpus* objetiva, principalmente, a explicação das etapas de construção e dos critérios postulados pelo(a) autor(a) para a realização da proposta. Como já debatido, a criação de *corpus* é, até o momento, escassa no contexto roraimense (e em todo o contexto amazônico). Por esse motivo, ou seja, por ser o *corpus* aqui debatido um pontapé inicial no estado, as fases metodológicas foram concebidas também de acordo com esse cenário, tendo para a sua execução completa a necessidade de dois momentos.

O primeiro momento se resumiu ao **reconhecimento da estrutura jornalística**, que se justificou pela ausência de discussões acerca da identidade da FBV, o que impossibilitava o conhecimento e a delimitação dos gêneros que poderiam ser considerados para as análises. Para tanto, o momento de reconhecimento da FBV, ocorrido entre setembro e novembro de 2023, foi realizado através de uma investigação dos jornais da Folha entre os anos de 2022 e 2017, nessa ordem decrescente – datas selecionadas para que o material a ser organizado fosse bastante atual. Nessa análise, definiu-se a observação da estrutura de um mês de publicação por ano. Como critério, na tentativa de contemplar o trabalho anual da FBV (início, meio e fim do ano), foi estipulado um intervalo de dois meses para cada análise, sendo examinados, então: 2022 –

dezembro, 2021 - fevereiro, 2020 - abril, 2019 - junho, 2018 - agosto e 2017 - outubro.

O objetivo central desse momento foi o de assimilar as colunas presentes no jornal, os gêneros de cada coluna e a frequência de cada um deles, o que fez com que a etapa seguinte, a da construção do *corpus*, fosse mais bem desenvolvida. Também nessa primeira fase se definiu que os gêneros das colunas que fossem de reprodução nacional e não de produção da FBV (mesmo com a sua diversidade) seriam dispensados da coleta, tendo em vista que um dos objetivos do trabalho era a organização de critérios para a construção das amostras de textos jornalísticos que atestassem, exclusivamente, a identidade da FBV. Nessa proposta, uma coluna foi dispensada - a Coluna Esplanada.

Por fim, ainda diante dessa etapa, concluiu-se que os gêneros mais presentes na FBV eram o artigo de opinião, a crônica, a notícia, a reportagem, o comentário opinativo, a legenda, o convite, o edital e o anúncio. Dentre eles, escolheram-se os que, juntos, atendiam a estes critérios: I - aproximações com as diretrizes do projeto Pró-norma; II - regularidade; e III - estrutura e função textuais.

O critério I se referiu ao projeto realizado em âmbito nacional, mencionado na Introdução deste artigo, que, dentre algumas direções, tem selecionado trabalhar com 7 gêneros textuais-discursivos jornalísticos e, para cada uma dessas amostras, tem reunido cerca de 20.000 palavras. Como esse projeto foi utilizado como modelo para este estudo, consideraram-se, inicialmente, as suas diretrizes para a escolha dos gêneros e para as fases da organização deste material. Contudo, em alguns momentos, foi indispensável ir além das diretrizes do Pró-norma, resultando na elaboração de critérios próprios. Por isso, visualiza-se abaixo a combinação de critérios empregados em toda a elaboração do *corpus*.

Quadro 1 - Diretrizes empregadas no *corpus*: combinações entre critérios do Pró-norma e aspectos específicos da FBV

CRITÉRIOS PRÓ-NORMA	CRITÉRIOS ESPECÍFICOS FBV
Delimitação dos gêneros textuais-discursivos (Proposta: 7 gêneros) <u>Critério adaptado → 5 gêneros</u>	Regularidade (Análise dos gêneros frequentes na FBV)
Delimitação de palavras (Proposta: 20.000 para cada gênero) <u>Critério adotado</u>	Estrutura e função textuais (Análise de especificidades dos gêneros presentes na FBV, com estrutura e função bem definidos)
Recorte temporal (Proposta: busca por textos recentes) <u>Critério adotado</u>	Presença de colunas específicas (Análise da relação entre colunas e gêneros)
Diversidade temática (Proposta: seleção de textos com temas variados) <u>Critério adotado</u>	
Textos originalmente do estado em análise (Proposta: busca por textos produzidos por indivíduos do estado) <u>Critério adaptado → descarte de colunas de reprodução nacional e, dentro do possível, seleção de textos produzidos por autores do estado</u>	

Fonte: Elaboração própria

Como demonstrado no quadro acima, a proposta do projeto Pró-norma pauta-se em realizar a coleta de 7 gêneros para a composição do *corpus* jornalístico, sendo eles: o artigo, a crônica, a notícia, o editorial, a carta de leitor, a entrevista e a tirinha – tendo sido esses gêneros, inicialmente, buscados na FBV. Desse total, precisou-se excluir aqueles que não atendiam aos critérios – no caso, o critério II – postulados por este trabalho especificamente para a Folha, como será detalhado a partir de agora.

No que se refere ao critério II – regularidade –, consideraram-se todos os gêneros que foram publicados com frequência de 2017 a 2022. Dos 7 gêneros

trabalhados no Pró-norma, 4 não tiveram regularidade na FBV, ou seja, foram descartados por descumprimento do critério II – foram eles: o editorial, a entrevista, a carta de leitor e a tirinha. A explicação que se dá para a não periodicidade na publicação de editoriais e cartas de leitor confirma o fato, apontado por Hellen e Alecrim (2013), de que gêneros opinativos são menos presentes na Folha. Quanto à entrevista, esse gênero foi praticamente nulo no material observado, pois o Grupo possui uma Rádio própria que busca se ocupar da realização do gênero em questão. Quanto às tirinhas, são publicadas, apenas, pela Coluna Esplanada, que não foi considerada para análise.

Sobre o critério III, o de estrutura e função textuais, notou-se a presença frequente de outros gêneros na FBV, representados por textos com estrutura e função comunicativa bem estabelecidas. Nesse sentido, incluíram-se no *corpus* os gêneros comentário opinativo e reportagem, que demonstram uma especificidade do jornal examinado.

Diante de todos os critérios, o *corpus* de Roraima não conseguiu seguir com a proposta inicial de conter 7 gêneros textuais-discursivos, pois a diversidade apresentada pelo jornal local não permitiu tal ação. Sendo assim, diminuiu-se de 7 para 5 gêneros, atendendo ainda à necessidade de substituir alguns deles através dos novos critérios (regularidade e estrutura e função textuais) e levando em consideração que muitos gêneros aparecem e desaparecem da FBV com muita facilidade. À vista disso, como se vê no quadro 2, o *corpus* do PRR escrito culto apresentou a seguinte configuração:

Quadro 2: Gêneros considerados para as amostras de textos jornalísticos roraimenses

	GÊNEROS	JUSTIFICATIVA
1	ARTIGO	Conforme Pró-norma + Regularidade + Estrutura e função
2	CRÔNICA	Conforme Pró-norma + Regularidade + Estrutura e função
3	NOTÍCIA	Conforme Pró-norma + Regularidade + Estrutura e função
4	+ REPORTAGEM	Regularidade + Estrutura e função
5	+ COMENTÁRIO OPINATIVO	Regularidade + Estrutura e função

Fonte: Elaboração própria

Esses 5 gêneros acontecem e se dividem em colunas específicas da FBV, sendo elas: coluna *Opinião* – com o aparecimento dos gêneros artigo e crônica –

, coluna *Social com Shirley Rodrigues* – especialmente com o gênero comentário opinativo –, e os demais gêneros (notícia e reportagem) coletados da estrutura geral da Folha, através das editorias de Cidade, de Política, de Polícia e de Variedades.

Com a delimitação dos gêneros concretizada, o segundo momento do trabalho se resumiu à **organização efetiva do material**. Para isso, mais uma vez a proposta apoiou-se nas diretrizes do projeto Pró-norma, estipulando a média de coleta de 20.000 palavras para cada gênero, sendo realizada a seleção de um texto por mês. Com esse objetivo, foi determinado como ponto de partida o último mês do ano de 2022, considerando, unicamente, o aspecto de recência. A proposta selecionou ainda um dia base para a coleta, o dia 21 de cada mês (escolhido por ser a data de implantação/aniversário da FBV). Nos casos em que o texto não fosse encontrado no dia estipulado, determinou-se que aconteceria o que se chamou de *recuo* diário. Assim, quando o texto, proposto para ser recolhido de acordo com os seus critérios, não foi encontrado no dia base (21), foi feito o recuo de dia por dia. Quando o texto foi encontrado, seguiu-se para o dia 21 do próximo mês a ser analisado. A motivação de cada recuo correspondia à especificidade de cada gênero, sendo, por vezes, explicada pela não publicação da coluna ou pela não publicação do gênero naquela edição.

Com a determinação dos passos a serem realizados com o *corpus*, a elaboração do material incluiu recorte e coleta, organizados em 4 fases (todas disponibilizadas no *corpus*), sendo elas: (i) **original** – imagem formato jpeg, da folha A4 completa em que o gênero estava inserido; (ii) **recorte** – imagem formato jpeg, do gênero recortado da folha A4; (iii) **texto editável** – imagem jpeg, transformada em docx.; e (iv) **versão final** – documento docx. revisado e editado pela autora, com centralização, orientação retrato, fonte *Times New Roman* - tam. 12, sinalização de título/subtítulo (negrito e sublinhado) e autor/coluna ou editoria/legenda (marcados pelos sinais de < >); e identificação do gênero e sua respectiva numeração (negrito e sublinhado).

É preciso ressaltar ainda que, mesmo nos arquivos finais, o *corpus* manteve a originalidade dos textos, permanecendo inclusive com as ocorrências de palavras sem espaçamento, erros de digitação e repetição de palavras (acontecimentos comuns em alguns textos), o que mais uma vez demonstra a preocupação em efetivar um *corpus* com representatividade local.

Todas as atividades aqui descritas com foco na elaboração do *corpus* viabilizaram o seguinte resultado:

Quadro 3: Informações gerais do *corpus*

GÊNERO		PALAVRAS	ARQUIVOS	PERÍODO
ARTIGO		20.629	29	dez-2022 a dez-2019
CRÔNICA		20.247	44	dez-2022 a set-2018
NOTÍCIA		20.017	72	dez-2022 a jan-2017
REPORTAGEM		20.088	36	dez-2022 a jan-2020
COMENTÁRIO OPINATIVO		20.239	52	dez-2022 a set-2018
Total:	5 gêneros	101.220 Palavras	233 arquivos	dez-2022 a jan-2017

Fonte: Elaboração própria

Em síntese, assim foi o trabalho realizado para a constituição do primeiro *corpus* para análises do português escrito culto de Roraima.

CONSIDERAÇÕES FINAIS

Tendo em vista a falta de bibliografia (sócio)linguística suficiente na região, os novos espaços, os novos trabalhos e as novas parcerias de pesquisa garantem uma maior autonomia para os estudos sobre o PRR, bem como um olhar mais amplo para o contexto que o cerca. Assim, com a união de pesquisadores voltados à criação de *corpora*, espera-se que, daqui para a frente, continuem a surgir novos trabalhos (sócio)linguísticos locais, envolvendo todas as modalidades e níveis de análise da língua, além de encorajar demais estudiosos a analisarem o cenário atual do estado para estabelecerem o que ainda pode ser efetuado na colaboração desse avanço científico.

Por meio da criação deste *corpus*, percebeu-se que elaborar um banco de dados sobre o português - no caso, escrito culto - roraimense, para além do amplo conhecimento linguístico necessário, requer um entendimento minucioso acerca da história política e geográfica do estado. Somente assim, foi possível alcançar autonomia e propriedade na elaboração de critérios específicos, etapa traçada, neste estudo, como a mais importante para que se pudesse assumir a representatividade do material como um instrumento de análise dos estudos (sócio)linguísticos do PRR.

O perfil do estado do extremo norte possibilitou o cenário de organizar os critérios de maneira a seguir algumas diretrizes nacionais, mas, ao mesmo tempo, oportunizou a liberdade de ajustar/manifestar as peculiaridades de acordo com o cenário local, o que favorece o fundamento de que é somente respeitando a sócio-história do estado que se obterá um material que alcance efetivamente os falantes dessa comunidade.

Esse ponto foi fortemente considerado porque se percebe que, pela história repleta de descontinuidades no jornalismo local (Correia, 2021), ainda há por parte dos jornais uma busca por identidade organizacional e estrutural (Gonçalves, 2024), o que exigiu uma análise sistemática sobre o jornalismo roraimense para que, então, fossem criados critérios específicos que verdadeiramente dialogassem com o cenário e seus respectivos dados linguísticos.

Por último, o crescimento no que se refere ao objetivo de criar *corpora* no estado - pois, afinal, com este, somam-se 4 *corpora* em andamento - inclui Roraima no caminho da compreensão de suas características linguísticas. Especialmente este material aqui construído, a ser disponibilizado em breve à comunidade acadêmica em um repositório preparado para recebê-lo, já possibilitará o andamento da descrição da(s) norma(s) que permeia(m) o PRR escrito culto. Essas futuras informações serão corroboradas pelos (sócio)linguistas que se debruçarem nas análises desses materiais.

REFERÊNCIAS

BAGNO, M. Cinquenta anos do NURC. *In*: OLIVEIRA Jr., M. (org.). NURC - 50 anos: 1969/2019. São Paulo: Parábola Editorial, 2019. p. 9-18.

BAGNO, M. **Gramática Pedagógica do Português Brasileiro**. São Paulo: Parábola Editorial, 2011.

BIAZOLLI, C. C. **Posição de clíticos pronominais em duas variedades do português**: interrelações de estilo, gênero, modalidade e norma. 2016. 381 f. Doutorado (Doutorado em Linguística e Língua Portuguesa) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, 2016.

BIAZOLLI, C. C.; BERLINCK, R. de A. Apresentação. *In*: BIAZOLLI, C. C.; BERLINCK, R. de A. (org.). **Gêneros textuais-discursivos no estudo de processos de variação e mudança**. Campinas: Pontes Editores, 2021a. p. 7-12.

BIAZOLLI, C. C.; BERLINCK, R. de A. Por que investigar processos de variação e mudança linguísticas por meio de gêneros textuais-discursivos? *In*: BIAZOLLI, C. C.; BERLINCK, R. de A. (org.). **Gêneros textuais-discursivos no estudo de processos de variação e mudança**. Campinas: Pontes Editores, 2021b. p. 13-38.

BIAZOLLI, C. C.; SENE, M. G.; AZEVEDO, L. O.; ZAMBRANO, P. C.; ALMEIDA, M. A. Gêneros textuais-discursivos do jornal *Folha de S.Paulo*: da construção da amostra a resultados preliminares. *In*: VIEIRA, S. R.; ROCHA, J. A. M. da; SILVA, L. F. A. da; LIMA, L. de C. de; LIMA, M. D. A. de O. (org.). **Varição linguística, ensino e interfaces**: resultados e propostas. São Paulo: Pimenta Cultural, 2023. p. 60-86.

CORREIA, C. M. **Jornalismo e memória a construção da política nos jornais de Roraima (1907-1988)**. 2021. 129 f. Dissertação (Mestrado em Comunicação Social) – Universidade Federal de Roraima, Boa Vista, 2021.

CRUZ, N. S.; PROCÓPIO, E. S. Léxico roraimense nas redes sociais: um estudo da conta @kabocando no Instagram. **Revista Falange Miúda**, v. 7, n. 2, p. 37-50, 2023.

FARACO, C. A. **Norma culta brasileira**: desatando alguns nós. São Paulo: Parábola, 2008.

FREITAG, R. M. K. NURC, um banco de dados sociolinguístico. *In*: OLIVEIRA Jr., M. (org.). **NURC - 50 anos: 1969/2019**. São Paulo: Parábola Editorial, 2019. p. 125-134.

FREITAS, A. **Geografia e História de Roraima**. 5 ed. Manaus: Belvedere, 1997.

GONÇALVES, L. C. S. **Amostras de textos jornalísticos roraimenses: da organização do material à identificação de características linguísticas**. São Carlos, 2024. (Relatório de Qualificação - Mestrado - apresentado ao Programa de Pós-Graduação em Linguística da Universidade Federal de São Carlos).

GONÇALVES, L. C. S. Um catálogo dialetal de São Luiz do Anauá. **Philologus**, v. 27, n. 81, p. 2126-2144, 2021.

HELLER, B.; ALECRIM, W. A Folha de Boa Vista e a desintrução na Terra Indígena Raposa Serra do Sol. **Revista FAMECOS - Porto Alegre**, v. 20, n. 1, pp. 211-229, jan./abr. 2013.

HUNSTON, S. **Corpora in applied linguistics**. Cambridge: Cambridge University Press, 2002.

LABOV, W. **Padrões sociolinguísticos**. São Paulo: Parábola Editorial, 2008 [1972].

LUZ, D. S. de B. **Reflexões sobre linguagem e identidade de maranhenses residentes em Boa Vista-RR**. 2013. 145 f. Dissertação (Mestrado em Letras) - Universidade Federal de Roraima, Boa Vista, 2013.

NASCIMENTO, S. **Estudo prévio da percepção do roraimense sobre o português de Roraima**. 2022. Monografia (Trabalho de Conclusão de Curso em Letras português-literatura) - Universidade Federal de Roraima, Boa Vista, 2022.

PIMENTEL, R. L. B. P. **O jornalista, o repórter e a mídia impressa em Roraima**. 1996. Monografia (Trabalho de Conclusão de Curso) - Universidade Federal de Roraima, Boa Vista, 1996.

PROCÓPIO, E. S. Constituição de um cópús para o português de roraima (SÉC. XVIII). **Revisa do GELNE**, Natal/RN, v.25, n. 3; e31584, setembro, 2023.

PROCÓPIO, E. S. Dicionário de palavras e expressões do português de Roraima. **Revista Philologus**, v. 27, n. 81, p. 960-968, 2021.

PROCÓPIO, E. S.; SILVA, A. C. A. Oitcenta - empréstimo e variação lexical no português de Roraima. **Revista Falange Miúda**, vol. 6, n. 2, p. 17-28, 2021.

PROCÓPIO, E. S.; SILVA, E. O. Neologismos no Português de Roraima. **Muiraquitã**, v. 10, n. 2, p. 246-267, 2022.

PROCÓPIO, E.; SILVA, P. L. da. Cruzamentos vocabulares com o topônimo 'Roraima'. **Estudos da Língua(gem)**, v. 21 n. 21, p. 9-32, jul-2023.

SARDINHA, T. B. **Linguística de Corpus**. 16. Ed. Barueri: Editora Manole, 2004.

SARDINHA, T. B. Linguística de Corpus: histórico e problemática. **D.E.L.T.A.**, v. 16, n. 2, p. 323-367, 2000.

SENE, M. G.; CHAVES, L. B.; PROCÓPIO, E. Evidências sociais do falar roraimense nas redes sociais: uma análise da página South BV City dos Memes. **Muiraquitã**, Rio Branco-AC, 2024 (no prelo).

SILVA, P. S. R.; VIEIRA, J. G. Uma breve análise histórica do jornal Folha de Boa Vista e suas influências políticas e ideológicas. **Norte Científico**, v.5, n.1, p. 143-159, 2010.

SOUZA, C. M. Considerações sobre a inserção social dos migrantes gaúchos em Roraima. **História Oral**, v. 9, n.1, p. 29-47, 2006.

STEFANOWITSCH, A. **Corpus linguistics: A guide to the methodology** (Textbooks in Language Sciences 7). Berlin: Language Science Press, 2020.

VALE, A. L. F. **Migração e territorialização** - as dimensões territoriais dos nordestinos em Boa Vista / RR. 2007. 268 f. Tese (Doutorado em Geografia) - Universidade Estadual Paulista, Presidente Prudente, 2007.

VIEIRA, S. R. (Coord.) **Projeto Pró-norma plural: do continuum fala-escrita para a norma-padrão**. (em andamento)



VIEIRA, S. R.; LIMA, M. D. A. de O. (org.). **Variação, gêneros textuais e ensino de Português: da norma culta à norma-padrão**. Rio de Janeiro: Letras UFRJ, 2019.

Recebido em 05-05-2024

Aprovado em 14-05-2024