



CRIAÇÃO DE UM BANCO DE DADOS LINGUÍSTICOS E CONTRIBUIÇÕES PARA O ENSINO DA LÍNGUA PORTUGUESA

THE CREATION OF A LINGUISTIC DATABASE AND CONTRIBUTIONS TO PORTUGUESE LANGUAGE TEACHING

Raimundo Gouveia da Silva

Professor do Instituto Federal de Educação, Ciência e Tecnologia do Acre (IFAC)
raimundo.silva@ifac.edu.br

Iandra Maria Weirich da Silva Coelho

Professora Titular do Instituto Federal de Educação, Ciência e Tecnologia do Amazonas (IFAM)
iandrawcoelho@gmail.com

RESUMO

Esse artigo apresenta uma proposta de criação de um Banco de Dados Linguísticos, composto por amostras da Reserva Extrativista Cazumbá-Iracema, localizada no Acre, e discute suas contribuições para o processo de ensino e aprendizagem da Língua Portuguesa. A metodologia de coleta e organização desse Banco de Dados ancora-se nos pressupostos teóricos dos padrões sociolinguísticos, nos fundamentos empíricos da Teoria da Variação e Mudança Linguística e na metodologia de coleta e manipulação de dados em sociolinguística. Os resultados evidenciam contribuições do uso dessa amostra para a criação de propostas de ensino, com foco na língua em uso, identificação dos fatores socioculturais que influenciam o surgimento e permanência da variação linguística e realização de pesquisas no âmbito das línguas naturais.

PALAVRAS-CHAVE: Ensino; Banco de Dados Linguísticos; Língua Portuguesa; Variação linguística.

1 INTRODUÇÃO

Esse artigo traz um recorte da dissertação intitulada “Variação linguística na língua portuguesa: uma proposta de ensino dos modos verbais com uso de banco de dados linguísticos”, realizada no Mestrado Profissional em Ensino Tecnológico, do Instituto Federal de Educação, Ciência e Tecnologia do Amazonas, com foco na apresentação de um Banco de Dados Linguísticos composto por amostras de fala do núcleo da Reserva Extrativista Cazumbá-Iracema, localizado no Acre, (doravante Banco Reci), que serviu como material de apoio e referência para criação de uma proposta de ensino e aprendizagem dos paradigmas verbais indicativo e subjuntivo da Língua Portuguesa e ênfase na variação linguística (ANTUNES, 2007).

Diante desse contexto, o presente artigo apresenta uma discussão sobre os elementos utilizados para a coleta e organização do Banco Reci, assim como reflexões sobre a importância



desse material, constituído em formato de um produto educacional, cujo objetivo é contribuir com o processo de ensino e aprendizagem da Língua Portuguesa. As discussões levam em conta a significância em abordar o contexto da variação linguística, seja no âmbito regional e/o nacional, e propor novas metodologias de ensino e aprendizagem da língua materna com foco no uso de Bancos de dados e recursos tecnológicos (ARAÚJO *et al*, 2014a; 2014b).

Nesse enquadramento, cabe destacar que no Brasil a compilação e uso de Banco de Dados com foco no processo de ensino e aprendizagem da língua materna ainda é embrionária, com poucos estudos que destacam tal relevância (SIMÕES; MELO, 2008; Monguilhott *et al.*, 2016). Dessa forma, buscamos evidenciar as principais contribuições e a necessidade em constituir amostras que podem servir como acervo para a criação e desenvolvimento de propostas de ensino, a partir do material sociolinguístico coletado, além de servir como “fonte privilegiada para a descrição do português brasileiro” (FREITAG; MARTINS; TAVARES, 2012, p. 918).

Para a coleta, metodologia e organização do Banco Reci, tomamos como base os padrões sociolinguísticos (LABOV, 2008), concepções relacionadas à sociolinguística variacionista (FREITAG; MARTINS; TAVARES, 2012), fatores que implicam na metodologia de coleta e manipulação de dados em sociolinguística (FREITAG, 2015), assim como aspectos técnicos na coleta de dados linguísticos orais (OLIVEIRA JUNIOR, 2014).

1.2 A PESQUISA SOCIOLINGUÍSTICA E A CRIAÇÃO DE BANCOS DE DADOS

Na pesquisa sociolinguística brasileira, o Banco de Dados é considerado a espinha dorsal, uma vez que por meio desse recurso podemos encontrar tendências de correlação entre as variáveis estudadas na comunidade pesquisada, ou realizar uma comparação com outras comunidades em virtude da padronização na metodologia da coleta de dados (FREITAG, 2015).

Um Banco de Dados é um conjunto de informações armazenadas sobre pessoas, lugares, de maneira estruturada. Esta organização tem por finalidade estabelecer o maior número possível de informação e trazer mais eficiência durante a realização de uma pesquisa. Nesse sentido, Simões e Melo (2008, p. 10), destacam que:

É justamente reconhecendo a importância de nossa diversidade cultural e linguística e os múltiplos olhares investigativos que ressaltamos a importância do Banco de Dados como fonte de pesquisa e memória disponibilizada.



Recurso imprescindível quando direcionamos nossas inquietações para a expansão da língua portuguesa e para a inclusão das diferentes comunidades.

Com base nas considerações de Simões e Melo (2008), justificamos a criação e uso de tais Bancos, no âmbito das pesquisas sociolinguísticas que, por meio de estudos empíricos sobre a heterogeneidade da fala em uma comunidade, tem por objeto de estudo os padrões de comportamento demonstrados através da fala e que podem contribuir para o surgimento e a manutenção do português brasileiro, uma vez que as “mudanças na língua podem assim, estar correlacionadas com mudanças na posição dos subgrupos com os quais o falante se identifica” (LABOV, 2008, p. 327).

Com isso, temos uma comunidade de fala que tem por princípio o agrupamento de pessoas que compartilham costumes, crenças, valores e, principalmente, relações interpessoais, independentemente dos traços linguísticos. Para Labov (2008), esse agrupamento de pessoas não compartilha necessariamente as mesmas variáveis linguísticas, mas dos mesmos valores dessas variáveis, sendo reconhecidos por seus pares como membros dessa comunidade.

A constituição de uma amostra que busca considerar tais variedades do português brasileiro leva em conta características acerca da seleção de seus informantes e cada projeto que constitui seu Banco de Dados busca atender a um conjunto de procedimentos metodológicos padronizados, a fim de possibilitar a realização de novos estudos e a descrição mais acurada da amostra que pode ser utilizada para a análise de diferentes fenômenos linguísticos.

Nesse sentido, no que diz respeito à constituição do Banco Reci, cabe destacar que o mesmo foi organizado de acordo com características sociodemográficas, enfatizando um padrão regular de estratificação socioeconômica das variáveis e as correlações entre tais variáveis linguísticas e categorias sociais primárias, como sexo, idade, escolaridade, entre outros. As características desse Banco seguem na próxima seção.

1.3 O USO DE BANCO DE DADOS LINGUÍSTICOS NA PRÁTICA DE SALA DE AULA

De acordo com Freitag, Martins e Tavares (2012, p. 918), “os Bancos de Dados Linguísticos de fala (especialmente os que seguem a orientação da sociolinguística variacionista) têm sido fonte privilegiada para a descrição do português brasileiro”. Na mesma direção, Monguilhott *et al.* (2016, p. 115) destacam que “os Bancos de Dados das variedades



do português, além de serem fonte para investigações sob a perspectiva da sociolinguística e suas interfaces, também se constituem em farto instrumento a ser utilizado nas aulas de Língua Portuguesa”.

Considerando a relevância das variedades linguísticas no processo de ensino e aprendizagem da língua materna, acreditamos que o professor pode utilizar estratégias metodológicas em sua prática pedagógica que valorizam a diversidade, tanto da comunidade em que a escola está inserida, como da fala na sala de aula, incentivando a construção de novas aprendizagens e valorizando os conhecimentos socioculturais e linguísticos dos alunos e dos moradores de uma determinada comunidade de fala.

É preciso, portanto, que a escola e todas as demais instituições voltadas para educação e a cultura abandonem esse mito da “unidade” do português no Brasil e passem a reconhecer a verdadeira diversidade linguística de nosso país para melhor planejarem suas políticas de ação junto à população amplamente marginalizada dos falantes das variedades não-padrão (BAGNO, 2007, p. 18).

Simões e Melo (2009, p. 20-21) também destacam a importância dos Bancos de Dados para o ensino da Língua Portuguesa, já que seu uso e pesquisa pode conduzir ao universo da diversidade cultural, contribuir para análise da grande quantidade de variantes próprias da nossa língua, assim como evidenciar “a riqueza dos diferentes falares de nossa cultura linguística”. Para os autores.

É justamente reconhecendo a importância de nossa diversidade cultural e linguística e os múltiplos olhares investigativos que ressaltamos a importância do Banco de Dados com fonte de pesquisa e memória disponibilizada. Recurso imprescindível quando direcionamos nossas inquietações para a expansão da Língua Portuguesa e para a inclusão das diferentes comunidades. (SIMÕES; MELO, 2009, p. 21).

De maneira geral, com relação à prática pedagógica, “alguns professores estão acostumados ao hábito de sistematização gramatical e utilizando modos ultrapassados de concepção de ensino de Língua Portuguesa” (BRITO; MATTOS; PISCIOTTA, 2003, p. 20). Essa discussão leva em conta a necessidade em trabalhar com diferentes variáveis sociais em sala de aula, tais como: faixa etária, sexo (masculino, feminino), escolaridade, *status* econômico, escolaridade etc., fatores inerentes às variáveis linguísticas (WEINREICH; LABOV; HERZOG, 2006), a fim de demonstrar a influência desses fatores que contribuem para o surgimento e a manutenção do modo de falar.

Dessa forma, quando o aluno tem consciência do seu modo de falar pode ter mais autonomia para apropriar-se das regras da norma culta e usá-las quando for necessário,



(CALLIAN; BOTELLHO, 2014), e isso pode partir desde o contexto linguístico dos alunos, possibilitando a partir dessa realidade linguística uma aprendizagem autônoma, desafiadora, prazerosa e significativa, contribuindo para a formação de cidadãos críticos, criativos, reflexivos e com capacidade comunicativa.

Nesse sentido, é importante ressaltar a relevância em trabalhar com ênfase na variação linguística, pois além de proporcionar um rico espaço de aprendizagem, é “inerente à língua”, e faz-se necessário, portanto, “discutir sobre ela em sala de aula” (PEREIRA; SILVA; SILVA, 2013, p. 108). Nessa perspectiva, considerando essa prática docente, ressaltamos a importância da criação de propostas de ensino e aprendizagem que tenham como ponto de partida o vernáculo natural dos estudantes, com o objetivo de compreender melhor os fatores que colaboram para o surgimento e a manutenção da linguagem. Para isso acontecer de forma satisfatória e autônoma é necessário que os “professores de Língua Portuguesa tenham conhecimento de que a linguística faz parte da vida de cada um e saibam como trabalhá-la dentro da sala de aula, de maneira a estimular o interesse de seus alunos” (PEREIRA; SILVA; SILVA, 2013, p. 111).

Dessa forma, faz-se necessário um trabalho diferenciado com uma abordagem mais reflexiva de se trabalhar com a língua, de forma que o aluno possa tornar-se um pesquisador de sua própria linguagem, identificando os elementos que influenciam seu falar cotidiano. Nesse sentido, o professor tem o papel de “fazer da sala de aula um laboratório de linguagem e atribuir aos alunos o papel de investigadores linguísticos” (GÖRSKI; COELHO, 2009, p. 84).

2. METODOLOGIA

Para a coleta, metodologia e organização do Banco Recí, tomamos como base os padrões sociolinguísticos (LABOV, 2008), concepções relacionadas à sociolinguística variacionista (FREITAG; MARTINS; TAVARES, 2012), fatores que implicam na metodologia de coleta e manipulação de dados em sociolinguística (FREITAG, 2015), assim como aspectos técnicos na coleta de dados linguísticos orais (OLIVEIRA JUNIOR, 2014).

A constituição do Banco Recí foi realizada por meio de entrevistas semiestruturadas. A amostra contou com entrevistas de 36 informantes, divididos em três células sociais, formando os critérios para a obtenção dos dados: sexo (masculino e feminino); idade (a partir dos 15 anos); escolaridade (Educação Básica e não alfabetizado).



A proposta de compilação desse Banco ancora-se nos pressupostos teóricos dos padrões sociolinguísticos (LABOV, 2008) e fundamentos empíricos da Teoria da Variação e Mudança Linguística (WEINREICH; LABOV; HERZOG, 2006). A coleta e organização dessa amostra estão delineadas na metodologia de coleta e manipulação de dados em sociolinguística (FREITAG, 2015).

A compilação do Banco Reci teve como propósito constituir uma amostra significativa para servir como fonte para a elaboração de uma proposta de ensino e aprendizagem da Língua Portuguesa, com foco na variação da língua e no uso de recursos tecnológicos para facilitar o trabalho docente.

A constituição dessa amostra se consolidou como um produto educacional que pode contribuir para além da elaboração de propostas de ensino e aprendizagem, sendo possível ampliar os Bancos de Dados Linguísticos referentes à Língua Portuguesa como fontes privilegiadas para a descrição da língua, servir como material para pesquisas em diferentes áreas e propiciar uma visão ampla das variedades linguísticas no âmbito nacional, com foco especialmente para a coleta e divulgação da fala dos povos tradicionais e para a constituição dos perfis sociolinguísticos dessas variedades que preconizam a língua, compreendida desde uma perspectiva variável e mutável.

Isso implica na possível contribuição para a ampliação dos Bancos de dados de fala independentes, especialmente da região norte¹, tais como o Bancos de Dados do Rio de Janeiro - PEUL-, o Banco de Dados VARSUL - Variação linguística na região Sul do Brasil e o Banco de Dados VALPB - Variação linguística no Estado da Paraíba, entre outros, proporcionando uma visão ampla das variedades do português que podem auxiliar em diferentes linhas de investigação, na perspectiva da sociolinguística variacionista.

Considerando tal propósito, a constituição do Banco Reci foi realizada por meio de entrevistas semiestruturadas. A amostra contou com entrevistas de 36 informantes, divididos

¹ Cabe ressaltar que no estado do Acre, em 1988, foi elaborado o Atlas Etnolinguístico do Acre – ALAC, através da elaboração de glossário do Vale do Acre. Em virtude da riqueza da pesquisa, os dados levantados, foram disponibilizados a outros pesquisadores e serviu de alicerce para o surgimento do Centro de Estudos Dialetológicos do Acre – CEDAC – objetivando armazenar Banco de Dados resultantes da pesquisa de doutorado, acrescido de outras coletas, com vistas a trabalhos variados no universo da linguagem, particularmente à feitura do Atlas Etnolinguístico do Acre – ALAC.



em três células sociais, formando os critérios para a obtenção dos dados: sexo (masculino e feminino); idade (a partir dos 15 anos); escolaridade (Educação Básica e não alfabetizado).

A proposta de compilação desse Banco ancora-se nos pressupostos teóricos dos padrões sociolinguísticos (LABOV, 2008) e fundamentos empíricos da Teoria da Variação e Mudança Linguística (WEINREICH; LABOV; HERZOG, 2006). A coleta e organização dessa amostra estão delineadas na metodologia de coleta e manipulação de dados em sociolinguística (FREITAG, 2015).

Selecionamos o núcleo da Reserva Extrativista Cazumbá-Iracema por acreditarmos que no meio rural encontra-se, provavelmente, uma das falas mais conservadoras, o que, por sua vez, pode servir de objeto de análise para outros pesquisadores, contribuindo, dessa forma, para a compreensão dos fatores socioculturais e de traços e perfis sociolinguísticos das variedades faladas nesta comunidade.

Entre as considerações avaliadas antes da coleta, ressaltamos a escolha da comunidade de fala, a determinação das variáveis da análise e o número de informantes. A amostra foi coletada a partir de três células sociais que formam os critérios da pesquisa e são categorias tradicionalmente consideradas pela sociolinguística variacionista: sexo, faixa etária, escolaridade, além de outras características gerais: ter nascido na comunidade, não ter se ausentado por mais de dois anos no período da adolescência e ser aceito pelos moradores do núcleo como membro da comunidade, como exposto no Quadro 1.

Quadro 1. Critérios socioculturais para a elaboração da entrevista

FATORES SOCIOCULTURAIS	
Sexo	Masculino e Feminino
Faixa etária	Faixa I - 12 informantes de (15 a 20 anos). Faixa II - 12 informantes de (25 a 40 anos). Faixa III - 12 informantes de (a partir dos 45 anos).
Continuação Escolaridade	Faixa I – Informantes cursando o Nível Médio na comunidade. Faixa II – Informantes com Ensino Fundamental II incompleto. Faixa III – Informantes com Ensino Fundamental I incompleto ou sem instrução escolar.



Características Gerais	Ter nascido na comunidade e não ter se ausentado por mais de 2 anos. Ter mais de 14 anos. Pais naturais da comunidade ou de outras comunidades, mas com mais de cinco anos morando na reserva. Intervalo de quatro anos entre as faixas etárias. Serem aceitos pelos outros pares como membro da comunidade de fala.
-------------------------------	--

Fonte: SILVA (2018, p. 43)

Para obtenção dos dados foram entrevistados 36 membros do núcleo da Reserva, estratificado em três grupos, por meio dos seguintes critérios: 12 informantes na faixa etária de 15 a 20 anos, cursando o nível médio na respectiva comunidade, preferencialmente, seis membros do sexo masculino e seis do sexo feminino. Para o segundo grupo de informantes, optamos por 12 membros do núcleo da reserva, na faixa etária de 25 a 40 anos, com o Ensino Fundamental II incompleto, sendo também dividido por sexo. E o terceiro, compreende 12 pessoas de ambos os sexos, a partir dos 45 anos. Nesse grupo, entrevistamos informantes sem instrução escolar e/ou pessoas com o Ensino fundamental I incompleto.

Os cuidados metodológicos adotados levam em conta que em uma pesquisa para criação de um Banco de Dados, os pesquisadores não podem organizar os informantes e elaborar os questionários apenas para alcançar seus objetivos, como afirmam Monguilhott *et al.* (2016, p. 114):

Os pesquisadores não realizam uma coleta propriamente para ser fonte de suas investigações particulares, mas os projetos em todo o país constituem seus Bancos de Dados considerando aspectos mais amplos justamente para dar conta de investigações nos diferentes níveis da língua e com diferentes abordagens da sociolinguística, considerando também suas interfaces.

Nessas circunstâncias, todas as pesquisas sociolinguísticas adotam padrões metodológicos que visam descobrir e elencar o máximo de informações possíveis de seus informantes, destacando os fatores internos e externos que colaboram para o surgimento e a permanência de uma comunidade de fala, já que são essas informações que podem ser objetivos de investigações por outros pesquisadores.

A escolha do perfil dos informantes teve como base alguns critérios, à luz da teoria laboviana. Foram selecionados falantes nascidos e residentes na comunidade do núcleo da



Reserva Extrativista Cazumbá-Iracema, com pais naturais também da comunidade, ou de outras comunidades, porém, com mais de cinco anos residindo na reserva, a fim de não perder a autenticidade e dificultar o aparecimento de outras influências nos resultados.

De acordo com a amostragem exemplificada, foi obtido um agrupamento de dezoito células sociais, considerando os procedimentos clássicos labovianos, Labov (2008), que determina a multiplicação entre os fatores das variáveis sociais. Assim, teríamos 2 sexos x 3 níveis de faixa etária x 3 níveis de escolaridade, totalizando 18 células sociais.

Vale ressaltar que a amostra ideal para essa estratificação seria de cinco indivíduos em cada célula social (MOLLICA; BRAGA, 2003), o que resultaria em um total de 90 informantes, levando em conta que “em uma pesquisa de cunho variacionista, o ideal é selecionar quatro ou cinco informantes em cada célula, para evitar, durante o momento da entrevista e constituição do *corpus* da pesquisa, um comportamento linguístico idiossincrático ou enviesado [...]”. (GUY; ZILLES, 2007 apud VITÓRIO, 2014, p. 54). Contudo, no Brasil, os Bancos de Dados costumam ser constituídos a partir de 2 a 3 informantes por célula social (TAVARES; MARTINS, 2014).

Embora a literatura aconselhe um número ideal para amostragem desse tipo de pesquisa, devido à escassez de tempo e recursos, tornou-se inviável a coleta de dados com esse número de indivíduos em cada célula, em virtude de diferentes obstáculos enfrentados, em especial as particularidades de acesso à essa comunidade de fala, já que o núcleo da Reserva Extrativista Cazumbá-Iracema está localizado em uma área de difícil acesso, principalmente nos meses de junho e julho, período em que o rio seca na região, evidenciado os problemas de caráter logístico (SANTOS, 2007). Outros aspectos fazem referência à necessidade de autorização junto ao ICMBio, o aceite dos gestores locais da comunidade do pesquisador que fará a coleta de dados e a documentação linguística e a dificuldade em encontrar informantes específicos².

Nesse sentido, visando minimizar tais obstáculos e estando em consonância com as diretrizes nacionais para o número de informantes, levamos em conta para o levantamento dos

² No caso da coleta mencionada, houve a dificuldade de encontrar informantes do sexo masculino de 15 a 20 anos que estivessem cursando o Ensino Médio, como os informantes do sexo feminino de 25 a 40 anos, com o Ensino Fundamental II incompleto.



dados, o número mínimo de informantes por célula, critério já amplamente adotado em Bancos de Dados como o PEUL, VARSUL e VALPB.

Além dos procedimentos metodológicos mencionados para o levantamento dos dados, também foi necessário: i) escolher o objeto de pesquisa; ii) verificar o número de habitantes do núcleo da reserva; iii) solicitar a autorização dos gestores da comunidade para a realização das entrevistas; iv) buscar pessoas que tivessem maior contato com os moradores para facilitar as entrevistas; e v) submeter o projeto ao Comitê de Ética e ao SISBio³.

Para a obtenção desses dados de fala, utilizamos a entrevista como estratégia de investigação. Apesar de inúmeras vantagens, também reconhecemos sua fragilidade, especialmente no contexto de estilo pergunta/resposta, uma vez que, de maneira geral, os informantes tendem a ficar se monitorando, não deixando fluir o seu verdadeiro vernáculo.

O questionário utilizado compreendeu 25 perguntas e a escolha das questões foi realizada a partir de elementos divididos em cinco seções, em que foi analisado o envolvimento dos informantes na comunidade, na família, redes sociais, infância, trabalho, buscando incentivar uma fala de maneira espontânea. Também foi incluída uma seção de perguntas referentes à produção linguística de base modal, a fim de verificar o uso dos modos indicativo e subjuntivo, com uso de alguns verbos, tais como: querer, esperar e acreditar, e expressões tais como, “lamento que” e o advérbio “talvez”.

Os dados foram gravados, organizados e transcritos, unicamente em uma linha, onde foi registrada a sintaxe real da fala dos informantes, considerando as hesitações, interrupções e demais fatores, além da ortografia oficial e os sinais de pontuação. Também foi realizada uma revisão das transcrições por meio de uma checagem manual dos áudios e comparação com as transcrições feitas, de acordo com os dados de fala e a realização de leituras e escutas de revisão,

³ Para a realização da coleta de dados no núcleo da Reserva Extrativista Cazumbá-Iracema submetemos o projeto para a apreciação do Comitê de ética. Além disso, solicitamos autorização ao Instituto Chico Mendes de Conservação da Biodiversidade- ICMBio. No momento da coleta dos dados na Reserva, torna-se imprescindível a apresentação dessa autorização, pois é ela que ratifica a liberação dos órgãos responsáveis pela fiscalização. Em virtude de o *lôcus* de pesquisa ser uma reserva extrativista e seus sujeitos povos tradicionais, fez-se necessário também submeter o projeto ao Ministério do Meio Ambiente, com um pedido de autorização para atividade com finalidade científica no sistema de autorização e informação em biodiversidade – SISBio.



para verificar se todos os dados transcritos estavam de acordo com as falas dos informantes e atribuir mais confiabilidade à transcrição dos dados.

3 DESENVOLVIMENTO

Ao concluir todo o processo, o Banco Reci foi organizado em formato de produto educacional e disponibilizado, por meio do repositório institucional do Programa de Pós-Graduação em Ensino Tecnológico. Para tanto foi compilado em dois formatos: *corpus* oral e escrito. A amostra oral possui mais de três horas de gravações, digitalizadas e armazenadas eletronicamente, e o *corpus* escrito possui um acervo de 30.798 palavras (constituído de amostras de fala, que foram gravadas, transcritas e, posteriormente, armazenadas eletronicamente)⁴. Essa divisão foi justificada pela necessidade de manter, no *corpus* oral, o anonimato dos informantes, com uma amostra de caráter restrito à biblioteca da instituição.

Vale ressaltar aqui, a importância do uso de repositórios institucionais para que os pesquisadores possam disponibilizar o acesso aos dados coletados e organizados em formato de produto educacional, de modo a impulsionar seu uso para o contexto de sala de aula, assim como a realização de estudos linguísticos que tratam sobre as variedades faladas da região norte do Brasil.

A fim de ampliar o acesso para que outros professores/pesquisadores tomassem conhecimento sobre o Banco de Dados Reci e suas principais características, além do repositório, também foi desenvolvido um *site*⁵, com o intuito de: i) auxiliar no trabalho pedagógico; ii) potencializar o ensino da língua materna no contexto da sala de aula; iii) fomentar novas formas de ensinar e aprender a língua materna; iv) servir como fonte para a elaboração de processos e produtos didáticos relacionados ao ensino e aprendizagem de fenômenos da Língua Portuguesa; v) contribuir para estudos descritivos da variação linguística e vi) fomentar a constituição e ampliação dos Bancos de Dados Sociolinguísticos, possibilitando novas pesquisas dessa natureza.

A criação desse *site* teve como principal intuito divulgar a pesquisa realizada e compartilhar informações que evidenciam a participação dos segmentos dos povos tradicionais na realidade sociolinguística brasileira, sob a constituição da realidade pluricultural da Língua

⁴ O Banco Reci (Corpus escrito) pode ser encontrado no repositório institucional: https://drive.google.com/file/d/1YekKfpwDUZ23VMZG-mOqE1PWZmUmOQ_w/view

⁵ Disponível em: <http://www.bancoreci.com.br/>



Portuguesa no Brasil. Contém a transcrição de algumas entrevistas, parte da proposta de ensino elaborada com auxílio da amostra, produções técnico-científicas oriundas da pesquisa, uma galeria de imagens, entre outros recursos. Para facilitar o acesso aos recursos do *site*, também foi criado um tutorial (SILVA; COELHO, 2018), que visa orientar o uso do Banco Reci e suas possibilidades para o contexto de ensino e aprendizagem⁶.

4 CONSIDERAÇÕES FINAIS

A compilação do Banco Reci, composto por dados linguísticos de populações tradicionais do núcleo da Reserva Extrativista Cazumbá-Iracema (BRASIL, 2007), teve como objetivo contribuir em diferentes aspectos: a) disponibilizar aos professores um conjunto de dados linguísticos reais como fonte para a elaboração de processos e produtos didáticos relacionados ao ensino e aprendizagem de fenômenos da língua portuguesa; ii) possibilitar o acesso dessa amostra para a comunidade científica que tenha interesse em pesquisar fenômenos relacionados à(s) variedade(s) sociolinguística(s) do norte do país; iii) contribuir para a ampliação dos Bancos informatizados de dados de fala dos povos tradicionais e o respeito às suas variedades linguísticas; iv) possibilitar a elaboração de novos projetos que possam incluir outros dados linguísticos coletados nas regiões do Acre, potencializando a coleta e estudos que visam à descrição do português falado nessa região e v) valorizar os conhecimentos socioculturais e linguísticos dos moradores e alunos dessa comunidade, enriquecendo o processo de ensino e aprendizagem.

Destacamos a potencial contribuição do Banco Reci no fomento às novas práticas para o ensino e aprendizagem da Língua Portuguesa, por meio de amostras reais de uso da língua com ênfase no estudo das variantes linguísticas, inserindo no contexto de sala de aula, diferentes possibilidades de ensino, de acordo com a realidade sociocultural e linguística dos alunos.

Além disso, destacamos também a relevante contribuição social, relacionada principalmente, com os possíveis aportes que as práticas de ensino dessa natureza pretendem alcançar em contextos de sala de aula das comunidades tradicionais. Trabalhar com dados

⁶ Disponível em: <https://drive.google.com/file/d/1w-P2yP5CkBsWNUCYi-CfY50KJEQGeuQF/view>.



autênticos desses povos leva em conta não apenas as ressalvas com relação à importância da variedade da nossa língua materna, mas também de aspectos como a valorização da região amazônica, especialmente, com relação à conservação dos recursos naturais, com destaque especial para a relação estreita que há entre esses povos e o ecossistema Amazônico e a vinculação da população que habita tradicionalmente esse diverso ecossistema.

Nesse sentido, acreditamos que a criação do Banco Reci possibilita evidenciar a realidade linguística brasileira e os aspectos sociais que a circunscrevem, permitindo evidenciar características importantes e dar a conhecer a localidade pesquisada a partir de suas falas, conhecer sobre sua história, a demarcação de territórios, o modo de lidar com os recursos naturais, entre outros fatores.

Acreditamos que essa amostra, compilada em formato de produto educacional, além de contribuir para o processo de ensino e aprendizagem, também pode ser significativa para outras áreas que discutem questões relacionadas ao uso da linguagem. Isso leva em conta aspectos relacionados ao contato entre as línguas, as variáveis sociolinguísticas, reflexões sobre a participação dos segmentos dos povos tradicionais na realidade sociolinguística brasileira, em relação a diferentes pontos de vista sob a constituição da realidade pluricultural da Língua Portuguesa no Brasil, a manutenção da identidade étnica desses povos, a perda linguística, fenômenos relacionados ao uso da língua materna, entre outros fatores, possibilitando trabalhos no domínio dos estudos linguísticos, assim como na área da Antropologia, da História, entre outros.

Por fim, sinalizamos esse contexto de pesquisa como desdobramento para trabalhos futuros que incluem a ampliação do Banco Reci, com dados linguísticos dos cinco polos que compõem a Reserva Extrativista Cazumbá-Iracema e de outras reservas extrativistas do Estado do Acre, com o objetivo de coletar e divulgar as variáveis linguísticas das reservas do Estado, possibilitando outras investigações na área do ensino, desde uma perspectiva sociolinguística, bem como auxiliar nos estudos morfológicos, fonológicos, sintáticos e/ou semânticos. Dessa forma, “a ampliação de banco de dados é uma tarefa que nunca se esgota, pois precisamos acompanhar a trajetória da variação e da comunicação linguística ao longo do tempo” (MONGUILHOTT *et al.*, 2016, p. 114).

REFERÊNCIAS



ANTUNES, I. **Muito além da gramática**: por um ensino de línguas sem pedras no caminho. São Paulo: Parábola, 2007.

ARAÚJO, A. S.; SANTOS, K. C. D.; FREITAG, R. M. K.. Redes sociais, variação linguística e polidez: procedimentos de coleta de dados. In FREITAG, R. M. K. (Org.), **Metodologia de coleta e manipulação de dados em sociolinguística**. São Paulo: Blucher, p. 99-116, 2014a.

ARAÚJO, S. S. F.; ALMEIDA, N. L. F. O projeto a língua portuguesa no semiárido baiano – fase 3: critérios de constituição e da amostragem do banco de dados. In: FREITAG, R. M. K. **Metodologia de coleta e manipulação de dados em sociolinguística**. São Paulo: Blucher, p. 27-47, 2014b.

BAGNO, M. **Preconceito linguístico o que é, como se faz**. 49ª ed., São Paulo: Loyola, 2007.

BRASIL. Ministério do Meio Ambiente. Dispõe sobre **Plano de manejo da Reserva Extrativista do Cazumbá**, 2007. Disponível em: <http://www.icmbio.gov.br/portal/images/stories/imgs-unidades-coservacao/resex_cazumba-iracema.pdf>. Acesso em: 5 mar. 2020.

BRITO, V. E.; DE MATTOS M. J.; PISCIOTTA, H. **PCNs de Língua Portuguesa**: a prática em sala de aula. São Paulo: Arte e Ciência, 2003.

CALLIAN, G. R.; BOTELHO, L. S. A análise linguística e o ensino de língua portuguesa: em busca do desenvolvimento da competência comunicativa. **Revista Eletrônica da Faculdade Metodista Granbery**, 16, p. 1-21, 2014. Disponível em: <<http://re.granbery.edu.br/artigos/NDk4.pdf>>. Acesso em: 5 mar. 2020.

FREITAG, R. M. K.; MARTINS, M. A.; TAVARES, M. A. Bancos de dados sociolinguísticos do português brasileiro e os estudos de terceira onda: potencialidades e limitações. **Alfa: Revista de Linguística**, 56(3), 2012.

FREITAG, R. M. K. Desafios teóricos-metodológicos da sociolinguística variacionista. In: PARREIRA, Maria Cristina et al. **Pesquisas em linguística no século XXI**: perspectivas e desafios teóricos-metodológicos. São Paulo: Cultura Acadêmica, p. 29-43, 2015. Disponível em: <<https://www.fclar.unesp.br/Home/Instituicao/Administracao/DivisaoTecnicaAcademica/ApoioEnsino/LaboratorioEditorial/serie-trilhas-linguisticas-n27.pdf>>. Acesso em: 10 set. 2020.

GÖRSKI, E. M.; COELHO, I. L. Variação linguística e ensino de gramática. **Working papers em linguística**, 10(1), 73-91, 2009.

OLIVEIRA JUNIOR, M. Aspectos técnicos na coleta de dados linguísticos orais. In: FREITAG, R. M. K. (Org.), **Metodologia de coleta e manipulação de dados em sociolinguística**. São Paulo: Blucher, p. 20-26, 2014.

LABOV, W. 2008. **Padrões sociolinguísticos**. São Paulo: Parábola Editorial, 392 p.

MOLLICA, M. C.; BRAGA, M. L. **Introdução à sociolinguística**: o tratamento da variação.



Editora Contexto, 2003.

MONGUILHOTT, I. O. S. et al. Metodologia de dados em escolas da rede pública e privada de ensino de Florianópolis. In: FREITAG, R. M. K.; SEVERO, C. G.; GÖRSKI, E. M. **Sociolinguística e política linguística: olhares contemporâneos**. São Paulo: Blucher, p. 113-136, 2016.

PEREIRA, B. G.; SILVA, J.; SILVA, J. B. M. da. Como trabalhar variação linguística e gramática em sala de aula: uma reflexão. **Revista Ribanceira**, 1(1), p. 107-119, 2013.

SANTOS, E. V. P. **Diálogos, Práticas e Espaços Participativos: A Participação da Comunidade da Reserva Extrativista Cazumbá-Iracema/Acre no Programa Biodiversidade Brasil-Itália**. Rio de Janeiro, RJ. Dissertação de Mestrado. Universidade Federal Rural do Rio de Janeiro, 139 f., 2007. Disponível em: <<https://tede.ufrj.br/jspui/bitstream/tede/632/1/2007%20-%20Edson%20Vanda%20Pereira%20dos%20Santos.pdf>>. Acesso em: 5 set. 2020.

SILVA, I. M. **As voltas que o modo dá: parâmetros funcionais da alternância indicativo/subjuntivo em espanhol**. Florianópolis, SC. Tese de doutorado. Universidade Federal de Santa Catarina – UFSC, 241 f., 2009. Disponível em: <<https://repositorio.ufsc.br/bitstream/handle/123456789/92650/266090.pdf?sequence=1>>. Acesso em: 25 set. 2020.

SILVA, R. G. da; COELHO, I. M. W da S. **Tutorial para o uso do banco de dados linguísticos RECI**. 2018. 28 f. Disponível em: <<http://ppget.ifam.edu.br/dissertacoes-defendidas/>>. Acesso em: 1 set. 2020.

SILVA, R. G. da; COELHO, I. M. W. da S. **Banco de dados linguísticos RECI: corpus Oral**, 2017, 1 DVD (218 min.43s) WAV.

SILVA, R. G.; COELHO, I. M. W. da S. Banco de Dados Linguísticos Reci – **Corpus escrito: 30.798 palavras**, 2017. DVD (PDF) Digital. Disponível em: <https://drive.google.com/file/d/1YekKfpwDUZ23VMZG-mOqE1PWZmUmOQ_w/view>. Acesso em: 1 set. 2020.

SILVA, R. G. **Variação linguística na língua portuguesa: uma proposta de ensino dos modos verbais com uso de banco de dados linguísticos**. 112 f. Dissertações. Mestrado Profissional em Ensino Tecnológico. Instituto Federal de Educação, Ciências e Tecnologia do Amazonas, 2018. Disponível em: <https://drive.google.com/file/d/1vSeVc_-VTvz4r7JH5mR8rsX8G8tHC22D/view>. Acesso em: 1 set. 2020.

SIMÕES, D.; MELO, E. M. A relevância dos bancos de dados para o ensino da Língua Portuguesa. **Revista ProLíngua**, 2(2), p. 12-24, 2008. Disponível em: <<http://www.periodicos.ufpb.br/index.php/prolingua/article/viewFile/13396/7604>>. Acesso em: 5 mar. 2018.

TAVARES, M. A.; MARTINS, M. A. O banco de dados fala-natal: uma agenda de trabalho. In: FREITAG, R. M. K. (Org.), **Metodologia de coleta e manipulação de dados em**



sociolinguística. São Paulo: Blucher, p. 71-78, 2014.

VITÓRIO, E. G. S.; LIMA A. A língua falada em alagoas: coleta e transcrição dos dados. In. FREITAG, R. M. (Org.). **Metodologia de coleta e manipulação de dados em sociolinguística.** São Paulo: Blucher, p. 49-60, 2014.

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

ABSTRACT

This paper shows a Linguistic Database creation idea, made by samples got from Cazumbá-Iracema Extrative Reserve, based in Acre and debate their contribution to Portuguese language teaching process. The Database collect and organization technique is grounded on theoretical assumptions of sociolinguistic patterns, in the empirical foundations of Theory of Variation and Linguistic Change and the data collection and manipulation methodology in sociolinguistics. The results show contributions of this sample to creation of teaching proposal, focused in Portuguese language, sociocultural identification factors that influence the arising and permanence of linguistic fluctuation and research of natural languages.

KEYWORDS: Teaching; Linguistic Database; Portuguese language; Linguistic variation.

Data de submissão: 2019

Data de aceite: 2020