

SOCIOLINGÜÍSTICA HISTÓRICA HISTORICAL SOCIOLINGUISTICS

Eliane da Rosa (UFRGS)¹

RESUMO: Desde a década de 80, muitos estudos de cunho histórico têm sido desenvolvidos até o presente momento sob uma nova perspectiva de investigação linguística. Este novo ramo da Linguística, chamado de Sociolinguística Histórica, tem contribuído de forma significativa para os estudos linguístico-históricos no mundo. Esta nova disciplina da Linguística tem por objetivo descrever e explicar fenômenos de variação e mudança linguística no passado por meio da utilização dos métodos de análise variacionista. O presente artigo pretende apresentar um breve relato sobre o percurso histórico desta nova linha de pesquisa, além de descrever seus pressupostos teórico-metodológicos.

PALAVRAS-CHAVE: Sociolinguística Histórica. Linguística Histórica. Língua Escrita.

ABSTRACT: Since the 80's, many historical studies have been developed up to now under a new perspective of linguistic investigation. This new field of Linguistics, called Historical Sociolinguistics, has contributed significantly to the historical studies in the world. This new discipline of Linguistics aims to describe and explain phenomena of linguistic variation and change in the past through the using of variationist methods of analysis. This article aims to present a brief description of the historical path of this new field of research, besides describing its theoretical-methodological presuppositions.

KEYWORDS: Historical Sociolinguistics. Historical Linguistics. Written Language.

Introdução

Partindo da premissa de que as línguas são sistemas complexos e dinâmicos que estão em constante evolução para suprir as necessidades comunicativas de seus falantes, a Sociolinguística Histórica tem por objetivo descrever e explicar os fenômenos de variação e mudança linguística em seu contexto social, cultural e histórico. Este campo da Linguística busca estudar os fenômenos linguísticos, que ocorreram no passado, através de amostras de língua escrita, pois em períodos mais antigos a única fonte de dados disponível ao linguista são os registros escritos.

A grande precursora da Sociolinguística Histórica foi Suzanne Romaine (2009[1982]), ao propor a união dos métodos de análise da Sociolinguística Variacionista com os da Linguística Histórica para investigar fenômenos linguísticos no passado. Com o advento desta

¹ Graduada em Letras-Licenciatura Plena em Língua Inglesa (ULBRA), especialista em Ensino-Aprendizagem em Língua Inglesa (Uniritter) e mestre em Morfologia e Fonologia (UFRGS). As pesquisas são voltadas para as mudanças linguísticas do português e do inglês. Aluna especial do Programa de Pós-Graduação em Letras da UFRGS. E-mail: elianedr19@gmail.com

nova proposta de investigação de cunho histórico, foi possível estudar processos de variação e mudança em tempos mais remotos, isto é, em uma determinada época do passado em que a tecnologia ainda não estava disponível para a gravação de fala.

Este artigo pretende apresentar um breve relato sobre o percurso histórico da Sociolinguística Histórica, assim como descrever seus pressupostos teórico-metodológicos.

1. Linguística Histórica

A Linguística Histórica é um ramo da Linguística que teve sua origem em fins do século XVIII num momento em que os estudos das línguas clássicas passaram a despertar o interesse dos eruditos pela origem e evolução das línguas. Esta área tem como objetivo investigar e descrever as mudanças pelas quais as línguas sofrem no eixo do tempo. Segundo Faraco (2005[1950]), este período é a marca cronológica do início de uma reflexão sistemática sobre as mudanças das línguas com base nos parâmetros da ciência moderna.

Nesta época, muitos estudiosos europeus começaram a se interessar por pesquisas a respeito do sânscrito, que é a língua clássica do povo hindu (Índia). Para a Linguística Histórica, o estudo do inglês William Jones, *The Sanscrit Language* (1786), é tido como o referencial dos estudos de cunho histórico. Este pesquisador, que atuava como juiz na cidade de Calcutá na Índia, observou que as raízes verbais e as formas gramaticais do sânscrito, do latim e do grego apresentavam muitas semelhanças estruturais. Com base nisso, Jones chegou à conclusão de que o sânscrito tinha um grau de parentesco com o grego e o latim. A partir desta descoberta, novos estudos sobre o sânscrito começaram a ser desenvolvidos desde então. Como resultado disso, surgiram diversas gramáticas e um dicionário da língua sânscrita. Além disso, em 1795, foi fundada uma Escola de Estudos Orientais em Paris em virtude do grande interesse pelo estudo das línguas clássicas. Esta escola tornou-se num influente centro de investigação histórica. Dentre os alunos que frequentaram esta instituição estavam Friedrich Schlegel e Franz Bopp, que, mais tarde, se tornariam nos precursores da gramática comparativa.

Em 1808, Schlegel publicou a obra *Über die Sprache und die weisheit der Inder* dando, assim, o pontapé inicial nos estudos comparativistas em solo alemão. Com seu estudo Schlegel conseguiu comprovar a hipótese de Jones sobre o parentesco entre o latim, o grego e o sânscrito, além de demonstrar que estas mesmas línguas apresentavam parentesco com o persa e a língua germânica. O autor concluiu ainda que essas semelhanças, isto é, que esse parentesco não era

por acaso, mas, sim, o fruto de influências mútuas, ou melhor, era a consequência de uma mesma origem. Todavia, para estabelecer o grau de parentesco entre as línguas era necessário compará-las entre si para, dessa maneira, chegar a sua ascendência comum.

No entanto, foi Franz Bopp quem deu o grande impulso aos estudos comparativistas ao publicar sua pesquisa intitulada *Über das conjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechischen, lateinischen, persischen, und germanischen Sprache* (1816). Por meio de seu estudo, Bopp conseguiu demonstrar que, pela comparação detalhada da morfologia verbal do latim, do grego, do sânscrito, do persa e do germânico, era possível determinar as correspondências sistemáticas que existia entre essas línguas. Com este procedimento, o estudioso deu origem ao método comparativo, que é a base metodológica primordial da Linguística Histórica. Este método tem por objetivo estabelecer o grau de parentesco entre as línguas.

O pressuposto de base é que entre elementos de línguas aparentadas existem correspondências sistemáticas (e não apenas aleatórias ou casuais) em termos de estrutura gramatical, correspondências estas passíveis de serem estabelecidas por meio duma cuidadosa comparação. Com isso, podemos não só explicitar o parentesco entre línguas (isto é, dizer se uma língua pertence ou não a uma determinada família), como também determinar, por inferência, características da língua ascendente comum de um certo conjunto de línguas. (FARACO, 2005[1950], p. 135)

Nos anos que se seguiram à publicação de seu livro de 1816, Bopp expandiu suas pesquisas comparativistas sobre outras línguas e, com os resultados obtidos o pesquisador publicou, em 1833, o trabalho que é considerado a obra básica da Linguística Histórica: *Vergleichende Grammatik des Sanskrit, Zend, Griechischen, Lateinschen, Litauischen, Gothischen und Deustchen*. Com este estudo, Bopp conseguiu descrever e elencar as principais subfamílias e famílias indo-europeias.

Nesse mesmo tempo, outro importante estudioso, John Grimm, buscou estabelecer as correspondências sistemáticas entre as línguas do ramo germânico. Como Bopp não levou em consideração o percurso histórico ao investigar o parentesco entre as línguas, Grimm optou pelo ramo germânico das línguas europeias numa sequência de 14 séculos, ou seja, por meio de sucessivos estágios de tempo.

Através deste novo procedimento, Grimm pôde observar e verificar que as línguas sofrem mutações, ou melhor, mudanças no decorrer do tempo. Sob a ótica deste novo prisma de análise linguística, o pesquisador acabou dando origem à *Gramática* ou *Linguística Histórico-comparativa*. Dessa forma, os estudos possibilitaram verificar que a sistematicidade das

correspondências entre as línguas tem uma estreita relação com o fluxo histórico e com a regularidade dos processos de mudanças linguísticas.

Com o advento do Estruturalismo no início do século XX, os estudos das línguas passaram por transformações metodológicas e o foco deixou de ser o estudo da evolução das línguas em favor do estado atual das mesmas.

O Estruturalismo teve seu marco inicial com a publicação de *Cours de Linguistique Générale* (1916)² de Ferdinand de Saussure, que é considerado o precursor desta nova abordagem. De acordo com a proposta de Saussure, a linguagem tem um lado individual e outro social que interagem mutuamente a ponto de não poderem ser concebidos um sem o outro. Para o linguista, o estudo da linguagem estaria dividido em duas partes: a psíquica (língua) e a psicofísica (fala).

Para o estruturalismo (SAUSSURE, 1916), a língua seria considerada o produto que o indivíduo registra passivamente, ou melhor, é a parte social da linguagem e exterior ao falante. A língua seria, portanto, composta por um sistema de signos, no qual existiria a união do sentido (significado) e da imagem acústica (significante). Enquanto a fala seria o ato individual de vontade e inteligência, a qual serviria como um meio para o falante realizar o código da língua e, dessa forma, exprimir seu pensamento pessoal e as combinações deste mesmo código linguístico (SAUSSURE, 1916).

É relevante destacar que na abordagem estruturalista, a língua passa a ser o objeto da Linguística em decorrência de esta fazer parte da coletividade e possuir um sistema homogêneo; ao passo que a fala não é considerada passível de análise por tratar-se apenas de manifestações momentâneas e individuais e por estar sujeita à influência de fatores externos, os quais, na maioria das vezes, não são linguísticos.

No que diz respeito ao estudo das línguas, Saussure sugeriu o estudo descritivo das línguas a partir de uma visão sincrônica ao invés de diacrônica. Em outras palavras, o linguista sugeriu que as investigações linguísticas fossem mais direcionadas ao estado atual da língua, que havia sido bastante negligenciado até então. Diante disso, o estudo das línguas ficou dividido em sincrônico (estudo de estados da língua) e diacrônico (estudo das mudanças da língua). Para o linguista suíço (COSTA, 2012, p. 118),

[...] o linguista deve estudar principalmente o sistema da língua, observando como se figuram as relações internas entre seus elementos em um determinado momento no tempo. Esse tipo de estudo é possível porque os falantes não têm informações acerca

² Esta obra, escrita por Saussure, foi publicada após sua morte. O material para publicação foi organizado por Charles Bally e Albert Sechehaye, junto com a colaboração de Albert Riedlinger.

da história de sua língua e não precisam ter informações etimológicas a respeito dos termos que utilizam no dia a dia: para os falantes, a realidade da língua é o seu estado sincrônico.

Dessa forma, com esta nova concepção metodológica, os estudos de cunho histórico, que utilizavam registros escritos como fontes de dados, foram sendo abandonados gradativamente em favor dos estudos baseados em amostras de língua falada. Em outras palavras, a sincronia, ou melhor, o estado atual da língua passou a ser o foco das pesquisas linguísticas. Era o início da preocupação com a descrição da língua falada que até então não havia recebido a devida atenção por parte dos pesquisadores, que estavam mais interessados em reconstruir o passado das línguas para descobrir qual era o ancestral em comum das mesmas. Mais tarde, com a implementação do Gerativismo e da Sociolinguística, o interesse pela investigação da língua falada se intensificou ainda mais. Consequentemente, as pesquisas de cunho histórico, ou melhor, a Linguística Histórica acabou perdendo espaço no âmbito acadêmico.

O Gerativismo surgiu com a publicação da obra *Syntactic Structures* (1957) de Noam Chomsky, como oposição ao Estruturalismo vigente até então. Este linguista propôs que a linguagem deveria ser considerada como um sistema de conhecimentos inatos, geneticamente determinado, inconsciente e modular, ou seja, a língua era formada por um sistema de regras. Segundo as premissas desta teoria, o falante possuiria dois tipos de gramáticas: a *universal* (GU) e a *particular* (GP). A Gramática Universal tratar-se-ia do estado inicial da linguagem, ou seja, quando a criança começaria a adquirir a língua, a qual está exposta. A evolução da GU resultaria na Gramática Particular (GP), que corresponderia às características próprias de cada língua. Desse modo, os elementos da língua fariam parte da GU, ao passo que a forma e a ordem como esses elementos são organizados, no sistema dessa língua, fariam parte da GP.

Para a Teoria Gerativa, a tarefa do pesquisador é estudar a competência linguística do falante e não o seu desempenho (uso real da língua), porque a capacidade para desenvolver a linguagem é uma habilidade inata do ser humano. Portanto, o interesse do Gerativismo é investigar o percurso psíquico da linguagem. Todavia, cabe ressaltar que, apesar de o foco das pesquisas terem se voltado para a língua falada, esta abordagem não leva em conta a fala produzida durante o uso real da língua, ou seja, a busca é por um “falante ideal”.

Diante do fato de a teoria chomskiana demonstrar desinteresse em inquirir sobre o desempenho do falante, um novo ramo da Linguística surgiu em oposição ao Gerativismo: a *Sociolinguística*. Esta, por sua vez, teve sua origem na década de 50 com o propósito de descrever e explicar a relação entre língua e sociedade, isto é, na linguagem chomskiana, o desempenho linguístico do falante. A língua passou, então, a ser vista como um sistema

heterogêneo passível de sofrer processos de variação e mudança linguísticas em virtude das influências exercidas por seus falantes. Quanto aos estudos de cunho histórico, estes permaneceram praticamente “esquecidos”.

2. Sociolinguística Histórica

Com a consolidação da Sociolinguística como campo de investigação linguística em meados da década de 60, um novo modelo teórico-metodológico do estudo da língua foi proposto por William Labov. Neste novo modelo de análise, chamado *Sociolinguística Variacionista* ou *Teoria da Variação*, os dados coletados passaram a ser submetidos a um tratamento estatístico de modo a possibilitar ao linguista a sistematização da variação existente na língua sob investigação. Como este novo método permitiu aos estudiosos ver a regularidade e a sistematicidade do caos linguístico, a Sociolinguística Variacionista acabou ganhando muitos adeptos. É importante destacar que, ao investigar processos de mudança em progresso, a Teoria da Variação, além de coletar dados em amostras de língua falada, também busca evidências no passado da língua para determinar de forma precisa se o fenômeno investigado está num estágio de mudança ou de variação.

Ao pesquisar fenômenos linguísticos em textos antigos, Suzanne Romaine notou que os registros escritos também apresentavam indícios de variação linguística da mesma maneira que a língua falada apresenta. Com base nessas observações, a linguista propôs a investigação de processos de variação e mudança linguística no passado com a utilização dos métodos de análise da Sociolinguística Variacionista. Em outras palavras, a estudiosa sugeriu a união dos métodos de análise da Teoria da Variação com os da Linguística Histórica para investigar fenômenos linguísticos em uma época em que não havia instrumentos para a gravação da fala.

A proposta da pesquisadora (ROMAINE, 2009[1982]) se fundamenta no fato de que a Sociolinguística não tem o privilégio exclusivo do estudo da variação, porque a Linguística Histórica e a Dialetoлогия se interessam pelo estudo da variação através do tempo e do espaço há muito mais tempo do que a Sociolinguística. A autora advoga ainda que, se a Linguística e a Sociolinguística têm o mesmo objetivo, que é estudar a língua, a Sociolinguística deveria ser testada com novos e diferentes tipos de dados. Isto quer dizer que o estudo sociolinguístico não deveria desenvolver pesquisas baseadas somente em amostras de língua falada, porque a

variação também ocorre na língua escrita, portanto esta, do mesmo modo, deveria ser considerada como fonte de dados. Segundo Romaine (op. cit., p. 13):

Anyone can observe that two samples of speech or writing are different. Sociolinguistic analysis can show us that these differences are objectively measurable, and that there are patterns in the choices which a speaker/writer does make, on the one hand, and can make, on the other. Since the choices are not entirely free, we need to observe the conditions or factors that may influence them; and it is here that sociolinguistic methodology is applicable.

Para a autora, a língua escrita deve ser tratada como fonte de dados da mesma forma que a língua falada, pois ambas são apenas formas diferentes de representar a mesma língua.

Durante séculos, a língua escrita serviu como objeto de estudo para os linguistas que buscavam descrever e explicar fenômenos linguísticos no passado e, principalmente, para obterem evidências acerca da fala de uma comunidade. No início do século XIX, foi a língua escrita que forneceu informações a respeito dos estágios evolutivos das línguas, portanto a língua escrita é uma importante fonte primária para as investigações linguísticas. Como diz Halliday (2006, p. 5), “it is part of the task of linguistics to describe texts; all the texts, including those, prose and verse, which fall within any definition of “literature”, are accessible to analysis by the existing methods of linguistics”. Isto quer dizer que o linguista deve estar apto para investigar e descrever fenômenos linguísticos tanto em amostras de língua falada quanto em amostras de língua escrita (incluindo não só correspondências, cartas e documentos pessoais, mas também textos literários).

Ao estudar os processos de variação do marcador relativo no Escocês Médio, Romaine conseguiu demonstrar que é possível investigar processos de variação e mudança em amostras de língua escrita por meio dos métodos de análise variacionista. O estudo da linguista consistia em examinar a relativização no Escocês Médio utilizando dados provenientes de uma amostra formada por diferentes tipos de textos escritos durante o reinado de James V. Para controlar as possíveis fontes irrelevantes de variação, Romaine limitou as amostras de textos cronologicamente (de 1530 a 1550³) e geograficamente (região de *Central Scots*).

A investigação de Romaine tinha como foco a variação entre as formas **WH** (*quhilk – which*), **TH** (*that*) e **Ø** (omissão do pronome relativo), que é uma característica dos textos pertencentes ao século 16. A análise dos dados coletados foi realizada com o uso do programa

³ Esta escolha de Romaine se baseou nas sugestões de McIntosh (1978, p. 42-43), que considera o período entre 1450 e 1550 o mais frutífero para fazer investigações linguísticas em virtude de esta época estar livre de tendências padronizantes no que se refere ao escocês médio. O autor também recomenda que é preciso restringir o período de tempo para menos de um século para evitar possíveis confusões entre os padrões cronológicos e regionais.

de regra variável *Cedergren-Sankoff* (CEDERGREN; SANKOFF, 1974), que estava em voga na década de 80. Com base nos resultados obtidos, a linguista pôde constatar que, no Escocês Médio, as estratégias de relativização das mulheres estavam mais próximas das normas vernaculares do que as estratégias dos homens, e, que a variabilidade estratificada estilisticamente poderia existir durante um longo período de tempo sem dar origem à mudança. Além disso, através deste trabalho, Romaine conseguiu testar e comprovar que os métodos de análise variacionista podem ser aplicados não só em dados provenientes da língua falada como em dados de língua escrita. Portanto,

Com a publicação da obra *Socio-Historical Linguistics: its status and methodology* (Romaine, op. cit.), as pesquisas de cunho histórico passaram a atrair novamente a atenção dos estudiosos. Como Romaine foi a primeira linguista a propor e a aplicar este novo procedimento metodológico em pesquisas de cunho histórico, a autora é considerada a precursora da Sociolinguística Histórica. Além disso, é necessário ressaltar que foi somente a partir da década de 90 que esta nova disciplina da Linguística começou a se firmar como campo de investigação científica. Desde então, diversos estudos têm sido desenvolvidos dentro da Sociolinguística Histórica em diversas partes do mundo (MONTGOMERY, 1995; SCHNEIDER, 2002; TUTEN, 2003; RAUMOLIN-BRUNBERG, 2005; MONTGOMERY, 2007; NURMI; NEVALA; PALANDER-COLLIN, 2009; NEVALAINEN, 2009; AYRES-BENNETT, 2014, entre outros). No que diz respeito à língua portuguesa, destacam-se os trabalhos de Lobo (2001), Carvalho (2003), Rumeu (2008) e Rosa (2015), por exemplo. É interessante destacar que, no Brasil, o número de trabalhos desenvolvidos, com base nos preceitos da Sociolinguística Histórica, vem aumentando gradativamente nos últimos anos.

3. Pressupostos Teórico-Metodológicos da Sociolinguística Histórica

Levando em consideração o fato de que as línguas sofrem alterações e mudanças em decorrência das influências exercidas por seus falantes e pelo meio em que estão inseridas, a Sociolinguística Histórica tem como meta descrever e explicar os fenômenos de variação e mudança linguísticas em seu contexto social, cultural e histórico. Conforme declara Hebda (2012, p. 238-239):

Given that the relationship of linguistic to extralinguistic variables falls within the scope of sociolinguistics, historical sociolinguistics must, by definition, investigate this same relationship from a diachronic perspective. In other words, while sociolinguistics inquires into the social uses of language (Chambers 1995/2003) in the sense of the linguistic choice people make, historical sociolinguistics is interested in what determined those choices in the past, the extralinguistic constraints involved being geographical background and mobility, gender, age, ethnic group membership, and the social network and socio-economic class of the speaker (Raumolin-Brunberg 1996).

Em linhas gerais, pode-se dizer que o procedimento de análise sociolinguística-histórica segue as mesmas etapas de uma análise sociolinguística: a) delimitação da variável dependente; b) descrição das variáveis independentes linguísticas e/ou extralinguísticas (que poderiam, por hipótese, estar influenciando a variável dependente); c) coleta de dados; d) transcrição e codificação dos dados coletados; e) quantificação dos dados por meio de um programa de regra variável. Porém, é relevante mencionar que, em decorrência da disponibilidade do registro escrito, o linguista sócio-histórico pode se defrontar com algumas restrições no que diz respeito à análise quantitativa dos dados. Este tópico será retomado mais adiante no decorrer desta seção.

Segundo Nevalainen e Raumolin-Brunberg (2012, p. 24), assim como em qualquer outro campo de investigação histórica, a Sociolinguística Histórica deriva sua *raison d'être* do *Princípio do Uniformitarismo*. Este princípio, formulado pela Geologia no século XIX, afirma que os fenômenos naturais que atuaram no passado são os mesmos que atuam, de maneira uniforme, na Terra nos dias atuais (LYELL, 1830). No campo da *Linguística*, este princípio foi introduzido por William Dwight Whitney (1867), exercendo grande influência na corrente neogramática.

Na Linguística Histórica, o Princípio do Uniformitarismo (WHITNEY, 1867; LABOV, 1994; JOSEPH; JANDA, 2003) passou a postular que as forças que operam para produzir a mudança linguística no presente são as mesmas que operaram no passado. Segundo Whitney (1867, p. 184):

The nature and uses of speech, and the forces which act upon it and produce its changes, cannot but have been essentially the same during all the periods of its history, amid all its changing circumstances, in all its varying phases; and there is no way in which its unknown past can be investigated, except by the careful study of its living present and its recorded past, and the extension and application to remote conditions of laws and principles deduced by that study.

O Princípio do Uniformitarismo é uma pré-condição necessária para que o linguista possa proceder à reconstrução histórica da língua que está sendo investigada. Dessa forma, ao utilizar o presente para explicar o passado, o pesquisador poderá inferir, através da observação

de processos em curso, como eram os fenômenos que operaram no passado. No entanto, ao propor o uso de métodos de análise variacionista com dados históricos, Romaine (op. cit.) sugere que é possível usar o passado para explicar o presente.

Para Romaine (op. cit.), Lass (2000) e Montgomery (2007), os registros escritos são compostos de formas ortográficas que podem estar sujeitas à interpretação linguística. Mesmo que o texto escrito não seja o reflexo simples e direto da fala, a escrita pode permitir a observação de traços gramaticais e de pronúncia considerando-se sua regularidade e aproximação com evidências de outras fontes, como comentários linguísticos contemporâneos e registros de períodos anteriores ou posteriores.

Romaine (op. cit., p. 16) assegura que é possível investigar e estudar a variação através da língua escrita, porque “the written hand shows regular variation in orthographic symbols which is conditioned by context in the same way that the spoken language displays allophonic variation”. Todavia, é necessário que o linguista saiba diferenciar o que pode ser considerado uma representação gráfica daquilo que pode ser uma representação fonético-fonológica do(s) segmento(s) de uma língua. Em outras palavras, é preciso que o estudioso esteja atento para não confundir o erro de um escriba com uma representação de um som. Para evitar este tipo de equívoco, Lass (op. cit.) recomenda que o pesquisador:

- 1^a. *seja falante nativo da língua investigada;*
- 2^a. *tenha conhecimento profundo da língua da época na qual o manuscrito foi escrito;*
- 3^a. *preste atenção às pistas e informações que o contexto do texto examinado fornece para responder às indagações da pesquisa em questão.*

O autor sugere ainda uma classificação para definir se o dado coletado do registro escrito é realmente significativo ou não para o estudo da variação e da mudança linguísticas. São estas as principais categorias dessa classificação (LASS, op. cit.):

- **Lixo** (*Garbage*): grafias que podem ser consideradas ‘*lapsus calami*’ ou um erro do escriba. Ex: *charope* em vez de *xarope*.
- **Variação puramente gráfica** (*Purely Graphic Variation*): grafias que apresentam certas características de uma tradição ortográfica ou da busca por traços etimológicos. Ex: consoantes geminadas → *bocca*.

- **Grafias fonologicamente significativas** (*Phonologically Significant Variation*): grafias que podem expressar fenômenos fonológicos, isto é, traços característicos da fala. Ex: *mininu* em vez de *menino*.

No que diz respeito ao *corpus* de análise, Romaine (op. cit.), Lass (op. cit.) e Montgomery (op. cit.) ressaltam que o pesquisador deve seguir alguns critérios para a composição do mesmo e que o caráter da análise, quantitativo ou qualitativo, será determinado pelo tipo de material disponível para a investigação. Em outras palavras, o linguista precisará identificar, descrever e selecionar os textos e os autores que produziram esses textos. Caberá ao linguista estabelecer seus próprios critérios quanto à seleção dos textos que comporão o *corpus*, isto é, decidir qual texto será incluído ou excluído da seleção e a justificativa para tal procedimento. Conforme Montgomery:

In corpus construction, texts are selected using accountable, transparent principles, e.g., they are the earliest or the most reliable edition of a printed work, or they form equal-sized text-samples chosen by time period, genre, type of author/speaker, locale, mode of language, etc. Validity and representativeness are multi-dimensional, ever-present issues in corpus construction regardless of the provenance or authorship/source of texts (MONTGOMERY, op. cit, p. 121).

Além disso, é interessante salientar que a composição do *corpus* é um trabalho que exige paciência, persistência e tempo em função de o linguista ter de selecionar o material de forma criteriosa e minuciosa. Assim, dependendo do tipo de fenômeno a ser investigado e do tipo de material disponível, este trabalho pode levar anos para ser concluído.

Montgomery (op. cit., p. 121-124) sugere ainda que o pesquisador leve em consideração algumas questões no momento de construir o *corpus*. Para facilitar esse trabalho, o autor estabelece algumas dimensões a serem levadas em conta. Dentre essas, destacam-se as dimensões: *textuais*; *temporais*; *sociais*; *espaciais*; e *de representatividade*.

a) **dimensões textuais**: o pesquisador deve verificar a natureza do registro escrito em si. Para isso, é preciso verificar que tipo de texto é, quão próximo da fala é o texto e qual é o estilo textual do manuscrito, porque os textos não são a representação simples e direta da fala. Em virtude disso é necessário ter cautela quanto a sua validade para a investigação histórica. Por exemplo, no caso dos textos literários (peças, diálogos ficcionais, etc.), estes podem apresentar variação exagerada e explorar estereótipos, dificultando, dessa forma, o trabalho do linguista em estabelecer a relação deles com os modelos da vida real.

b) **dimensões temporais:** o pesquisador deve buscar informações referentes a data em que o texto foi produzido e qual período da história da língua o texto representa. Neste caso, há uma grande vantagem em utilizar cartas, documentos, jornais, etc., como fontes primárias, porque é possível obter a data e o local de origem dos mesmos. Já com textos literários (peças, poesias, etc.) tem-se a desvantagem de ter apenas a data de publicação, pois o texto pode ter sido produzido muitos anos antes da publicação. Em consequência disso, torna-se difícil definir a que época pertence a língua retratada pelo autor.

c) **dimensões sociais:** o pesquisador deve tentar descobrir por quem o texto foi produzido, que informações pode obter sobre a vida daquele autor, e de quem a língua do texto representa. No que tange aos textos publicados, há grande desvantagem em precisar a identidade e o perfil social dos envolvidos (editor, impressor, etc.) na produção do texto por serem, geralmente, desconhecidos. Ao passo que há grande vantagem ao lidar com manuscritos pessoais, como cartas e bilhetes, pois permitem, na maioria dos casos, delinear o perfil sociocultural daqueles que produziram estes textos. Neste caso, o linguista pode obter estas informações através dos próprios manuscritos desses autores ou por intermédio de arquivistas ou historiadores.

d) **dimensões espaciais:** o pesquisador precisa saber a que região aquele texto pertence; qual é a nacionalidade ou a região de origem do autor, e se o texto, ou melhor, a variedade da língua é localizável, isto é, se pelo sistema ortográfico daquele autor é possível definir de que região ele provém. Em tempos mais antigos, era comum saber, por meio da escrita, o local de onde provinha um indivíduo, pois cada lugar tinha sua própria característica gráfica em decorrência da falta de uniformização no sistema ortográfico das línguas. Neste item é importante salientar que o estudioso pode buscar informações pessoais do(s) autor(es) em igrejas, cartórios, arquivos públicos, etc., por serem instituições que registram e preservam documentos.

e) **dimensões de representatividade:** o pesquisador precisa detectar de quantos indivíduos o texto provém, a que classe social esses indivíduos pertencem e quão generalizável são os padrões da língua que esses indivíduos evidenciam em sua escrita. Neste caso, a obtenção dessas informações dependerá da disponibilidade ou da sobrevivência dos manuscritos. Na maioria das vezes, o material que está à disposição da pesquisa histórica pode estar fragmentado, transcrito para uma linguagem atual, ter sofrido alterações gráficas ou, ainda, falsificações.

Para Montgomery (2007) e Schneider (2002), há uma grande diferença entre o planejamento de uma pesquisa sociolinguística em sincronia atual e uma pesquisa histórica, porque os sociolinguistas podem definir a amostra e os informantes, ao passo que a investigação

de cunho histórico sempre estará restrita à disponibilidade dos registros escritos. Em virtude disso, é comum surgirem certos problemas ao longo do desenvolvimento da análise sociolinguística-histórica. Dentre esses problemas, destacam-se:

- *a representatividade da amostra, que geralmente é pequena;*
- *a validade dos textos escritos para análise, que depende do autor (sua vontade em representar as formas da fala ou sua habilidade para tal);*
- *escolher entre uma análise quantitativa ou qualitativa;*
- *determinar os parâmetros de contextos extralinguísticos, que, muitas vezes, não é possível devido à falta de acesso às informações a respeito do escritor;*
- *a validade sócio-histórica, ou seja, a reconstrução da informação social necessária para interpretar os padrões de variação nos registros escritos nem sempre é simples e direta.*

De acordo com Kretzschmar e Stenroos (2012, p. 113-114):

A corpus of texts can never represent, in the sense of a randomized sample, a “real” historical population of speakers. In most periods, literacy has been confined to an elite minority, and the text types produced would represent a small part of the overall range of language use. In addition, the survival of texts is unpredictable but also not random: certain types of texts are more likely to have survived than others [...] For historical surveys dealing with written data, the choice of linguistic features and variables will depend on the realities of text production, the types of text produced, and the degree of institutionalization involved in the period and community surveyed.

Contudo estas dificuldades enfrentadas pelo pesquisador não impedem a execução da análise, que pode ser de caráter qualitativo, a qual, mesmo não oferecendo amplas generalizações, pode fornecer observações razoáveis sobre o fenômeno (MONTGOMERY, 1995). Como diz Labov (1994, p.11), é preciso fazer o melhor uso desses “dados ruins”, pois os documentos históricos sobrevivem por acaso, sem propósito, e a seleção que está disponível para pesquisa é o produto de uma série imprevisível de acidentes históricos. Romaine (op. cit., p. 122) e Joseph e Janda (2003, p. 14) também afirmam que, apesar de ser fragmentado e incompleto, o dado histórico é válido para a investigação histórica. Schneider (op. cit., p. 90-91) ressalta, ainda, que “working with written data requires somewhat more judgment and assessment than an analysis of audio recordings, but the difference is a matter of degree: essentially, with both approaches the goal is the same, and the pathways to reach it are very similar”.

Considerações Finais

As pesquisas desenvolvidas sobre o prisma da Sociolinguística Histórica têm contribuído de forma significativa para o enriquecimento dos estudos linguístico-históricos no Brasil e no mundo; pois a união dos métodos de análise variacionista com os da Linguística Histórica revolucionaram a forma de investigar fenômenos de variação e mudança linguística no passado de forma mais precisa e eficaz, além de permitir uma maior confiabilidade nos resultados das análises.

A partir da proposta de Romaine (2009[1982]), as pesquisas de cunho histórico voltaram a fazer parte do contexto acadêmico e a receber o devido respeito por ter a função de auxiliar os pesquisadores, que investigam fenômenos no presente, a entender a atuação e a propagação de determinados processos linguísticos desde os tempos mais remotos até os dias atuais. Com o advento da Sociolinguística Histórica, foi possível demonstrar e comprovar que a língua escrita também pode ser considerada como fonte de dados assim como a língua falada, porque ambas são apenas formas diferentes de representar a mesma língua.

Referências

AYRES-BENNETT, Wendy. From l'usage to le bon usage and back: norms and usage in seventeenthcentury France. In: RUTTEN, Gijsbert; VOSTERS, Rik; VANDENBUSSCHE, Wim (eds.). **Norms and usage in language history, 1600–1900: a sociolinguistic and comparative perspective**. Amsterdam & Philadelphia: John Benjamins, 2014, p. 173–200.

BOPP, Franz. **Über das Conjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprache**. Frankfurt-am-Main: in der Andreäischen Buchhandlung, 1816.

_____. **Vergleichende grammatik des sanskrit, zend, griechischen, lateinischen, litthauischen, gothischen und deutschen**. Berlin: Dümmler, 1833.

CARVALHO, Maria José. The transition from early to modern portuguese: an approach from historical sociolinguistics. In: BLAKE, Barry J.; BURRIDGE, Kate; TAYLOR, Jo (eds.). **Historical linguistics 2001**. John Benjamins Publishing, 2003, p. 59-69.

CEDERGREN, H.; SANKOFF, D. Variable rules: performance as a statistical reflection of competence. **Language** 50, 1974, p. 333-55.

CHOMSKY, Noam. **Syntactic Structures**. The Hague: Mouton, 1957.

COSTA, Marcos Antonio. Estruturalismo. In: MARTELOTTA, Mario Eduardo (Org.). **Manual de Linguística**. 2. ed. São Paulo: Contexto, 2012.

FARACO, Carlos Alberto. **Linguística histórica**: uma introdução ao estudo da história das línguas. São Paulo: Parábola Editorial, 2005[1950].

FITZMAURICE, Susan; SMITH, Jeremy. Evidence for the history of English: introduction. In: NEVALAINEN, Terttu; TRAUOGOTT, Elizabeth (eds.). **The oxford handbook of the history of english**. New York: Oxford University Press, 2012. p. 19-36.

HEBDA, Anna. Phonological variables. In: HÉRNANDEZ-CAMPOY, Juan M.; CONDE-SILVESTRE, J. Camilo (eds.). **The handbook of historical sociolinguistics**. Malden/Oxford: Wiley-Blackwell, 2012, p. 237-252.

HALE, Mark. **Historical linguistics**: theory and method. Malden: Blackwell Publishing, 2007.

HALLIDAY, M. A. K. **Linguistic studies of text and discourse**. v. 2. London/New York: A & C Black, 2006.

JONES, William. **The sanscrit language**: collected works. vol. 3. 1786, p. 34-5.

JOSEPH, Brian D; JANDA, Richard D. (eds.). **The handbook of historical linguistics**. Oxford: Blackwell Publishing, 2003.

_____. **Historical linguistics and sociolinguistics**: strange bedfellows or natural friends? 2011. Disponível em: <http://www.ling.ohio-state.edu/~ddurian/AWAC/Joseph_2011.pdf>. Acessado em: 17 jun. 2014.

KRETZSCHMAR JR., William A.; STENROOS, Merja. Evidence from surveys and atlases in the history of English language. In: NEVALAINEN, Terttu; TRAUOGOTT, Elizabeth (eds.). **The oxford handbook of the history of english**. New York: Oxford University Press, 2012. p.111-122.

LABOV, William. **Padrões sociolinguísticos**. Tradução de Marcos Bagno et al. São Paulo: Parábola Editorial, 2008 [1972].

_____. **Principles of linguistic change**: internal factors. v. 1. Cambridge: Blackwell, 1994.

LASS, Roger. **Historical linguistics and language change**. Cambridge: Cambridge University Press, 2000.

LOBO, Tânia Conceição Freire. **Para uma sociolinguística histórica do português no Brasil**: edição filológica e análise linguística de cartas particulares do Recôncavo da Bahia, século XIX. 4 Volumes. Tese (Doutorado em Filologia e Língua Portuguesa) – Universidade de São Paulo, São Paulo, 2001.

LYELL, Charles. **Principles of geology**. v. 1. London: John Murray, 1830.

MCENERY, Tony; WILSON, Andrew. **Corpus linguistics**: an introduction. 2. ed. Edinburgh: Edinburgh University Press, 2001[1996].

MACINTOSH, a. The dialectology of Medieval Scots: some possible approaches to its study. **Scottish Literary Journal**. Supplement n. 6, 1978, p. 38-44.

MONTGOMERY, Michael. The Linguistic value of Ulster Emigrant Letters. *Ulster Folklife*, n. 41, p. 1-16, 1995.

_____. Variation and historical linguistics. In: BAYLEY, Robert; LUCAS, Ceil. **Sociolinguistic variation**: theories, methods, and applications. Cambridge: Cambridge University Press, 2007. p. 110-132.

NEVALAINEN, Terttu. **Grasshoppers and blind beetles**: caregiver language in early modern English correspondence. In: NURMI, A.; NEVALA, M.; PALANDER-COLLIN, M. (eds). 2009, p. 137-164.

NEVALAINEN, Terttu; RAUMOLIN-BRUNBERG, Helena. Historical sociolinguistics: origins, motivations, and paradigms. In: HÉRNANDEZ-CAMPOY, Juan M.; CONDE-SILVESTRE, J. Camilo (eds.). **The handbook of historical sociolinguistics**. Malden/Oxford: Wiley-Blackwell, 2012, p. 22-40.

NURMI, A.; NEVALA, M.; PALANDER-COLLIN, M. (eds.). **The language of daily life in England (1400-1800)**. Amsterdam: John Benjamins, 2009.

PENKE, Martina; ROSENBACH, Anette. What counts as evidence in linguistics? Introduction. In: PENKE, Martina; ROSENBACH, Anette (eds.) **What count as evidence in linguistics**: the case of innateness. Amsterdam/Philadelphia: John Benjamins Publishing, 2007, p. 01-54.

RAUMOLIN-BRUNBERG, Helena. The diffusion of subject YOU: A case study in historical sociolinguistics. In: **Language Variation and Change**. Cambridge: Cambridge University Press, 2005, p. 55 – 73.

ROMAINE, Suzanne. **Socio-historical linguistics**: its status and methodology. Cambridge: Cambridge University Press, 2009 [1982].

RUMEU, Márcia Cristina de Brito. **A implementação do ‘você’ no português brasileiro oitocentista e novecentista**: um estudo de painel. 928 f. Tese (Doutorado em Letras Vernáculas) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2008.

ROSA, Eliane da. As vogais médias átonas finais no português brasileiro do século XIX. **Revista Galega de Filloxía**, v. 16, 2015, p. 147-167.

SAUSSURE, Ferdinand de. **Cours de linguistique générale**. Paris: Payot, 1916.

SCHALLEY, Andrea C. **Practical theories and empirical practice**: a linguistic perspective. Amsterdam/Philadelphia: John Benjamins Publishing, 2012.

SCHLEGEL, F. *Über die Sprache und die Weisheit der Inder*. Trad, francesa: **De la langue et de la philosophic des indiens**. Genebra: Mauget & Cherbuliez, 1809 [1808].

SCHENDL, Herbert. **Historical linguistics**. Oxford: Oxford University Press, 2009.

SCHNEIDER, Edgar W. Investigating variation and change in written documents. In: CHAMBERS, J. K.; TRUDGILL, Peter; SCHILLING-ESTES, Natalie (eds.). **The handbook of language variation and change**. Malden: Blackwell Publishing, 2002. p. 67-93.

STEPHENSON, Edward A. On the interpretation of occasional spellings. **American Dialect Society**, Alabama, n. 48, p. 33–50, 1967.

TUTEN, D. N. **Koineization in medieval spanish**. Berlin: Mouton de Gruyter, 2003.

WHITNEY, William Dwight. **Language and the study of language**: twelve lectures on the principles of linguistic science. New York: Charles Scribner & Company, 1867.