

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE LETRAS
DEPARTAMENTO DE LINGUÍSTICA
CURSO DE BACHARELADO EM LETRAS PORTUGUÊS-ALEMÃO**

CAROLINA FALCK GRIMM

**USO DE FORMAS ETIMOLOGIZADAS COMO CRITÉRIO DE GRAU DE
CULTURA DE REDADORES OITOCENTISTAS DO PORTUGUÊS BRASILEIRO**

Porto Alegre
2018

CAROLINA FALCK GRIMM

**USO DE FORMAS ETIMOLOGIZADAS COMO CRITÉRIO DE GRAU DE
CULTURA DE REDADORES OITOCENTISTAS DO PORTUGUÊS BRASILEIRO**

TCC: Trabalho de Conclusão de Curso apresentado ao Departamento de Linguística do Instituto de Letras da Universidade Federal do Rio Grande do Sul, como parte dos requisitos para obtenção do título de Bacharel em Letras; Universidade Federal do Rio Grande do Sul; Linguística Histórica

Orientadora: Prof^a Dr^a Valéria Neto de Oliveira Monaretto

Porto Alegre
2018

CAROLINA FALCK GRIMM

**USO DE FORMAS ETIMOLOGIZADAS COMO CRITÉRIO DE GRAU DE
CULTURA DE REDADORES OITOCENTISTAS DO PORTUGUÊS
BRASILEIRO**

TCC: Trabalho de Conclusão de Curso
apresentado ao Departamento de
Linguística do Instituto de Letras da
Universidade Federal do Rio Grande do
Sul, como parte dos requisitos para
obtenção do título de Bacharel em Letras;
Universidade Federal do Rio Grande do
Sul; Linguística Histórica

Orientadora: Prof^a Dr^a Valéria Neto de
Oliveira Monaretto

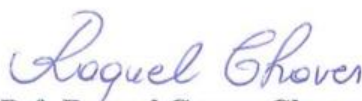
APROVADA: Porto Alegre, 20 de dezembro de 2018

Profa. Valéria Neto de Oliveira Monaretto
Dra. pela Pontifícia Universidade Católica do Rio Grande do Sul
Orientadora

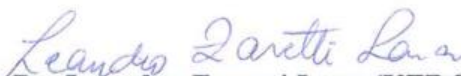
BANCA EXAMINADORA



Prof^a Dr^a Valéria Neto de Oliveira Monaretto (UFRGS)



Prof^a. Dr^a. Raquel Gomes Chaves (UFRGS)



Prof. Dr. Leandro Zanetti Lara (UFRGS)

RESUMO

Busca-se no presente trabalho aperfeiçoar um método de classificação objetiva de se aferir o grau de cultura de redatores oitocentistas brasileiros, proposto por Barbosa (2005), e testar se o mesmo pode ser utilizado em diferentes regiões do País. Para isso, foi calculado o uso de etimologias, como uma marca de uma norma culta escrita do século XIX, contrastado com sua taxa de acerto/erro, em cartas pessoais. O corpus analisado corresponde a cartas pessoais entre familiares de Júlio de Castilhos, personalidade importante da história social e política do estado do Rio Grande do Sul no século XIX, e cartas encaminhadas para Cícero Dantas Martins, Barão de Jeremoabo, da região do sertanejo baiano. Os dados do corpus dessa pesquisa são confrontados com os resultados de uma análise de uma amostra examinada por Barbosa (2005). Apesar de não ter sido confirmada a possibilidade de se aplicar a metodologia em cartas de outras regiões, a metodologia ora reexaminada tornou-se, preliminarmente, mais objetiva.

Palavras-chave: Metodologia em Linguística Histórica; Aferição de Norma Culta Século XIX; Etimologias.

ABSTRACT

The objective of this work is to improve a meta-objective method of measuring the degree of culture of nineteenth-century Brazilian writers, proposed by Barbosa (2005), and test whether the same method can be utilized in different regions of the country. In order to do so, the use of etymologies was calculated, as a mark of a written standard language from the nineteenth century, contrasted with their rate of right/wrong use, in personal letters. The analyzed corpus corresponds to a series of personal letters exchanged between relatives of Júlio de Castilhos, an important persona of the social and political history of the State of Rio Grande do Sul in the nineteenth century, as well as letters forwarded to Cícero Dantas Martins, Baron of Jeremoabo, from the countryside of Bahia. The corpus data of this research are confronted against the results of an analysis of a sample examined by Barbosa (2005). Although it was not possible to confirm the possibility of applying the methodology for letters from other regions, the methodology here re-examined became, preliminarily, more objective.

LISTA DE TABELAS

Tabela 1 – Relação de Formas Etimológicas em Cartas dos Familiares de Júlio de Castilhos (<i>corpus</i> 1).....	46
Tabela 2 – Relação de Formas Etimológicas dos Avós Ottoni por Barbosa (2005)	48
Tabela 3 – Relação de Formas Etimológicas em Cartas de sertanejos baianos a um Barão (<i>corpus</i> 2).....	49

LISTA DE QUADROS

Quadro 1 – Classificação de Graus de Erudição dos <i>Corpora</i> Analisados	51
---	----

SUMÁRIO

1 INTRODUÇÃO.....	8
2 A CONSTRUÇÃO DE UMA NORMA LINGUÍSTICA NO PORTUGUÊS BRASILEIRO	11
2.1 PORTUGUÊS BRASILEIRO: ORIGENS E DESENVOLVIMENTO	11
2.2 ESTUDOS SOBRE A NORMA LINGUÍSTICA DO PORTUGUÊS BRASILEIRO DO SÉCULO XIX	15
2.2.1 A crise sobre uma Norma Linguística do Português Brasileiro por Pagotto (1998, 2011)	15
2.2.2 Outros estudos sobre o perfil social do redator de sincronias passadas	29
3 DIRETRIZES DA PESQUISA	38
3.1 OBJETIVOS E HIPÓTESES	38
3.2 METODOLOGIA	42
3.3 AMOSTRA	44
4 RESULTADOS	46
5 CONCLUSÃO	53

1 INTRODUÇÃO

Este trabalho é fruto de pesquisas realizadas em atividades de Iniciação Científica sob orientação da professora Valéria Monaretto. O Projeto de Pesquisa “PARA UMA HISTÓRIA DO PORTUGUÊS BRASILEIRO GAÚCHO: variação e mudança fonológica em textos escritos oitocentistas” objetiva a construção e a análise de corpora de textos oitocentistas escritos no Rio Grande do Sul. Esse Projeto formou um banco de dados de documentos escritos por gaúchos no século XIX, do qual utilizamos uma parte como corpus para essa monografia, como cartas pessoais trocadas entre os familiares de Júlio de Castilhos, personalidade de grande importância histórica e política do estado.

Ambos trabalhos se inserem no campo da pesquisa de Linguística Histórica. De acordo com Faraco (2005), essa área de estudos propõe analisar e compreender o processo de mudança linguística. As línguas humanas não constituem realidades estáticas, portanto, estão em constante mudança.

Essa mudança é facilmente percebida por qualquer leigo que se deparar diante de algum material do século XIX, pois a língua escrita dessa época é diferente em alguns aspectos do momento atual, como a grafia de palavras, por exemplo. Textos escritos no passado são, pois, de difícil compreensão por causa de diferenças entre estágios de língua.

A mudança que estava em curso em determinada época dificilmente é percebida por seus falantes, já que normalmente não têm consciência de que a língua é dinâmica (FARACO, 2005, p. 14). A mudança, que é lenta e contínua, atinge apenas partes da língua e não, um todo. Eis um dos motivos pela não percepção de sua mobilidade por seus usuários.

A mudança é precedida por variações, que, em um certo momento atual da língua, podem de fato ser percebidas pelo falante em contraste com alguém de idade diferente, por exemplo. É importante salientar, que essas variações necessariamente não resultarão numa mudança. Como o próprio autor ressalta “[...] nem toda variação implica mudança, mas [...] toda mudança pressupõe variação [...]” (FARACO, 2005, p, 23).

Outro contraste que nos permite identificar eventuais mudanças em progresso é o contraste de língua falada e língua escrita. Normalmente a língua escrita é mais conservadora do que a oral, revelando certas incongruências do que é usado na fala. Contudo, é através da língua escrita que podemos examinar o passado das línguas.

Esse conservadorismo se dá, principalmente, pelo fato de que atividades escritas estão, em geral, ligadas a contextos sociais marcados de formalidade. Dessa forma, algumas inovações presentes na língua falada podem não ser aceitas de imediato na escrita. Mesmo assim, existem sim formas de estudarmos a língua no passado.

Como veremos ao decorrer do trabalho, o gênero textual mais utilizado para esses estudos são justamente cartas. Isso porque existe uma grande formalidade em torno da língua escrita, e cartas, quando trocadas entre conhecidos próximos, pessoas de mesma família, amigos em geral, são escritas de forma menos monitorada. E é justamente isso que o pesquisador interessado em estudar a língua do passado, precisa.

A língua menos monitorada será mais natural, portanto, mais vernacular. Será nesse cenário que possíveis variações, que ainda estão em conflito com outras, possam aparecer, em forma de indícios e vestígios. Portanto, é de suma importância que tenhamos em mãos documentos escritos em que os redatores estão mais à vontade e não se preocupem tanto com uma possível norma escrita da época.

Exemplificado isso, vamos aos fatos: não é simples encontrarmos esses materiais ideais. Parte dos materiais disponíveis são documentos extremamente formais oficiais ou cartas de redatores desconhecidos. Mesmo casos não tão extremos, como o nosso, apresentam complicações, pois corpora de grandes personalidades normalmente são compostos por pessoas escrevem com certo monitoramento.

Com isso, vimos a necessidade de achar algum facilitador para pesquisas históricas. É praticamente impossível olharmos para uma carta do século XIX e dizermos se ela é uma carta formal ou informal. Inclusive a primeira impressão que teremos é de que se tratam todas de cartas formais, pois fogem do que entendemos por formal atualmente. Ou seja, é preciso algum método para se identificar isso.

Para sabermos se algum documento de época antiga se trata de um texto formal ou informal, é preciso algumas informações sobre os redatores. Essas informações também são, em geral, escassas. A principal informação para esse caso, que seria interessante de se conseguir, é sobre os conhecimentos de língua daquele redator. Isso é tratado no presente trabalho como grau de cultura, uma vez que não temos como julgar o que é ser escolarizado e não escolarizado em uma época, em que muitos poucos escreviam e liam, como no caso do século XIX no Brasil.

Para cumprir com esse objetivo, situamos o contexto da Língua desde seu surgimento até sua chegada no Brasil, e todos os conflitos em torno do que era considerado norma culta para época. Esses capítulos são extensos pelo fato de ser impossível simplesmente pegarmos uma gramática e julgar essa o padrão de norma escrita seguido por todos. Esse fato já foi constatado pela pesquisa de Barbosa (2005), que conclui que os textos modelos do século XIX são, na verdade, os jornais da época, e que esses apresentavam o que parecia ser um marcador de prestígio pra época.

Com isso, Barbosa (2005) pressupôs corretamente que a utilização de etimologizações era um marcador de norma culta escrita para a época. Foi ele quem desenvolveu a ideia de que quanto mais se usasse etimologizações, maior seria o grau de cultura daquele redator. Uma vez que essa era uma característica de prestígio, ocorriam as ditas falsas etimologizações, ou seja, pessoas que supunham que aquela palavra descendia, por exemplo, do Latim, e, dessa forma, grafava-a erroneamente. Resumidamente trata-se de calcular a taxa de uso das etimologizações e contrastar com a taxa de acerto. Algumas etimologizações que encontramos em nossos *corpora*, por exemplo, são: *exceção*; *solenne*; *fructa*; *telegraphar*, etc.

Após essas seções sobre a história da nossa Língua, citamos alguns trabalhos de pesquisa de outros autores que buscaram o mesmo objetivo que este trabalho, que é o identificar traços dos perfis sociais de determinados redatores. Por fim chegamos, de fato, ao material estudado aqui por nós: o corpus de Júlio de Castilhos, que é contrastado com o corpus retirado do Corpus Eletrônico de Documentos Históricos do Sertão, com base na metodologia proposta por Barbosa (2005).

Em nossa análise de resultados, tentamos levar em conta todos esses fatores comentados aqui, o que se mostrou, na verdade, extremamente complicado. Isso se mostrou difícil, pois temos, como o objetivo desse trabalho, não só testar e tentar desenvolver melhor a proposta de Barbosa (2005), mas também checar a possibilidade dessa metodologia em redatores de outras regiões do Brasil. Dessa forma, nos deparamos com muito mais fenômenos que o imaginado. Talvez, justamente por isso, conseguimos abrir portas para novas ideias a serem desenvolvidas.

2 A CONSTRUÇÃO DE UMA NORMA LINGUÍSTICA NO PORTUGUÊS BRASILEIRO

Nessa seção abordaremos um pouco sobre a Língua que aqui estudamos, o português. Veremos tanto sua história, ou seja, sua origem e as mudanças pelas quais a Língua passou até se tornar o que era no século XIX, como também como a percepção que tinham dela nessa época. Com percepção queremos dizer as discussões em torno do que era a Língua e o que ela deveria ser, ou seja, as tentativas de normatização do Português Brasileiro.

2.1 PORTUGUÊS BRASILEIRO: ORIGENS E DESENVOLVIMENTO

Até os dias de hoje predominam crenças populares como nossa Língua “não ser correta” por não se assemelhar tanto com a fonética portuguesa europeia, ou como os falantes nativos brasileiros não “sabem falar português”. Isso tudo não é por acaso, mas sim resultado tanto da nossa colonização, feita por quem também ainda estava com sua língua em plena evolução linguística, como da futura tentativa de se criar, depois de anos de desenvolvimento da Língua portuguesa em território brasileiro, uma norma que buscasse uma similaridade com o português europeu (doravante PE).

A Língua Portuguesa pertence ao grupo de línguas ditas “românicas” ou “neolatinas”, que têm como ponto de partida o Latim. De leste para oeste da Europa temos o romeno, da Romênia, ou Rumânia, nos Balcãs; o italiano, na Itália; o francês, na França; na Península Ibérica ou hispânica, o castelhano, ou espanhol da Espanha e o português de Portugal. Todas essas línguas são o resultado da evolução do Latim, que se implantara numa vasta região da Europa decorrente de conquistas militares e do consequente domínio cultural e político de Roma a partir do século III a.C. (CÂMARA JR., 1975, p. 14-15). Não é surpresa, então, que encontraremos nos nossos *corpora* empréstimos justamente vindos do francês e do italiano.

A região de Portugal, que é parte a relevante para o nosso trabalho, teve ao que tudo indica, certas distinções étnico-culturais básicas das populações nativas atendidas pela tripartição administrativa romana, ou seja, assim como em outras regiões da Península Ibérica, recebeu influências externas variadas de povos que se instalaram antes dos romanos, como os celtas, visigodos, etc. Sendo assim, todo o território que corresponde hoje a Portugal devia apresentar aspectos étnicos, culturais e linguísticos próprios. Portanto, o Latim, pouco a pouco, radicou-se nesse território, e fez desaparecer as línguas nativas, e esse latim, já em forma de romance, passou a ser a língua nativa de toda a região.

Por volta do século VIII, o Latim propriamente dito já estava passando para outro estado de língua. Apesar de, com a ocupação islâmica na Península Ibérica ter ocorrido o domínio oficial da língua árabe, ainda houve disseminação da língua latina. Esse latim, já muito distante de suas origens e diversificado regionalmente, continuou a evoluir “na boca” das populações submetidas. Por conta do trabalho da Reconquista cristã, esse latim, dito moçárábico foi se encontrando com o Latim dos Cristãos.

É apenas no século XI, com a separação do reino de Leão e Castela, na zona atlântica, que surge o condado de Portugal e, portanto, a primeira língua nacional portuguesa: o romance peculiar (CÂMARA JR. 1975, p. 19). No século XIII, o português já apresentava uma língua literária *stricto sensu* em face do castelhano e do catalão, porém a língua escrita corrente e a prosa literária surgiram tardiamente e tiveram que substituir o hábito da redação em Latim.

O autor diz que é a língua escrita ser o que “reflete as condições gerais da língua comum nacional e acompanha a sua evolução.” (CÂMARA JR., 1975, p. 20) trata-se de um pressuposto. Mas que é por conta desse pensamento que se costuma dividir o português em um período arcaico, que vai até o século XV, e em um período posterior, que poderia ser chamado de moderno. Essas duas fases são distinguidas por traços característicos de natureza fonológica, gramatical e lexical bem demarcados. Além desses períodos, é citado pelo autor, um período clássico para os séculos XVI e XVII e outro pós-clássico, para os séculos subsequentes. O período clássico contém diferenças gramaticais nítidas, que, a partir do século XVIII, recebe mais a influência do francês na escrita.

A Língua Portuguesa desenvolveu-se, obviamente, de forma diferente quando trazida para o Brasil. O País, até o século XVI, povoado por índios, passou a ser também ocupado pelos portugueses e espanhóis, principalmente. Por conta dessa colonização, vieram muitos negros escravizados, trazidos da África. Portanto, essas foram as bases étnicas e culturais da formação da população brasileira, e é a partir dessa mistura que se desenvolverá a nossa Língua.

A Língua Portuguesa de Portugal chegou tanto sob a forma dos seus dialetos setentrionais como dos meridionais. Aparentemente houve certo equilíbrio na proporção entre a imigração do norte e a do sul. Câmara Jr. (*op. cit.*) acredita que, desse modo, no Brasil, houve, desde o início, condições novas para uma vida linguística particular e para o desenvolvimento de uma subnorma na língua comum, em face do português europeu. Essas novas condições ora se ressentiram e ora se mantinham mais nos padrões escritos do PE,

como é notado por um poema épico ainda no século XVI, encontrado em Pernambuco (CÂMARA JR., 1975, p. 31).

O cenário brasileiro de meados do século XVI era, portanto, constituído pelos “colonos” de origem portuguesa, falantes do Português Europeu, porém com claros traços específicos que foram se acentuando ao longo do tempo; e pela população indígena, africana e mestiça, que aprendeu o português de forma imperfeita. Além disso, tínhamos, ao lado do português, a *língua geral*, que se trata do tupi, principal língua indígena das regiões costeiras, porém, um tupi simplificado, gramaticalizado pelos jesuítas, portanto, transformado numa língua comum (TEYSSIER, 1997, p. 94)

Porém, na segunda metade do século XVIII, essa língua geral entra em decadência decorrente, entre diversos outros motivos, da chegada de muitos imigrantes portugueses atrás de minas de ouro e diamantes recém descobertas. É nesse mesmo século que as primeiras alusões a traços específicos do PB são documentadas, inclusive as diferenças da língua oral brasileira e da portuguesa, como dizer *minino* por *menino* (TEYSSIER, 1997, p. 95).

É apenas com a Independência, já no século XIX, que surge a questão nacional de se criar uma identidade própria. De acordo com Teyssier (1997), muitos brasileiros achavam impossível termos uma nação original, com cultura e literatura próprias, sem termos uma língua original. Essa questão preocupou principalmente escritores e filólogos, e, como o autor mesmo citou, gerou a polêmica da redação do Código Civil Brasileiro, texto muito analisado por Pagotto (2011), que será discutido mais extensamente na próxima seção.

Para nosso trabalho, cabe mencionar mais um fato relevante trazido por Teyssier (1997). A questão de que, apesar das tentativas de elaborações de mapas dialetais do Brasil, a nossa divisão é mais sociocultural do que geográfica:

As diferenças na maneira de falar são maiores, num determinado lugar, entre um homem culto e o vizinho analfabeto que entre dois brasileiros do mesmo nível cultural originários de duas regiões distantes uma da outra. (TEYSSIER, 1997, p. 98)

Sendo assim o autor divide o “brasileiro” em uma série de níveis. No topo, temos a língua das pessoas cultas, com variação entre o registro oficial estrito e o registro familiar livre; seguido pela língua vulgar das camadas urbanas gradativamente menos instruídas; e, por fim, os falares regionais e rurais. A importância desse fato para nosso trabalho será melhor elaborada na explicação do porquê da escolha dos nossos *corpora*.

O Brasil ficou, então, até há pouco tempo, sem ter uma gramática normativa bem estabelecida, o que provocou reflexos que vemos até os dias de hoje. De acordo com Bueno (1958), foi apenas a partir de 1911 que começamos a ter alguma interferência com objetivo de enquadrá-la dentro de normas científicas, seja por parte do governo ou de academias. É importante para nosso trabalho ressaltar o fato de que, no período galego-português, a prosa encontrava-se desnorteada, dado o fato dos copistas, tabeliães e cronistas tentarem constantemente aproximarem-se do latim.

Verifica-se, na língua escrita do português brasileiro do século XIX, o uso do *h* frequentemente sem necessidade alguma; *y* sem ter diferenciação alguma com o *i*; geminações de todas as espécies; e por fim, as falsas etimologizações, como *thesoura*, pelo fato talvez de existir a palavra *therouso*; *occeano* por conta da palavra *occidente* ou *phythysica* sob a influência de *physica* (BUENO, 1958, p. 289) são exemplos do grande reflexo que o latim teve na formação da nossa língua. Mesmo nos tempos clássicos, o latim ainda serviu de norma ortográfica, mesmo nos casos em que a consoante já houvesse evoluído foneticamente.

Por fim, levando em consideração o que foi abordado nessa seção, é importante ressaltar o que diz Câmara Jr.: as diferenças da língua padrão do Brasil e de Portugal não devem ser explicadas por um suposto resquício do tupi ou uma suposta profunda influência africana, mas sim, essencialmente do fato de a língua encontrar-se em dois territórios nacionais distintos e separadas. (CAMARA JR., 1975, p. 32)

A língua evoluiu de formas diferentes em Portugal e no Brasil, e muitas vezes sem coincidirem nas mudanças, mesmo tendo relações de vida social e uma cultura muito estreitas. Temos como exemplo a evolução da fonologia, que aqui no Brasil resultou do contato entre variados dialetos ultramarinos e a da língua padrão, enquanto em Portugal houve os fenômenos em *allegro* e grande insistência na sílaba tônica (*op. cit.*, p. 32).

Podemos notar, por fim, o porquê de certas influências no PB, e também, o motivo de tantas diferenças entre o PB e o PE. Essa quantidade de discrepâncias faz-se necessário ser evidenciada, pois ajuda a entender as dificuldades que os gramáticos e filólogos tiveram ao tentar normatizar a nossa Língua com base em certos aspectos do português lusitano, luta que será explicada melhor na próxima seção.

2.2 ESTUDOS SOBRE A NORMA LINGUÍSTICA DO PORTUGUÊS BRASILEIRO DO SÉCULO XIX

A Sociolinguística Histórica enfrenta muitos problemas ao querer examinar a variação linguística em textos escritos no passado. Dentre alguns, podemos citar a dúvida que se tem de se saber se quem escreve é culto ou não. Se o redator domina a língua escrita dentro de um padrão de época; se a linguagem de seu texto representa um vernáculo de época; de que época é o texto, e se quem o escreveu, seria um jovem ou não.

Abordaremos nas seções que seguem uma discussão sobre norma linguística brasileira, que se iniciou no final do século XIX, por meio de alguns estudiosos contemporâneos. Pretendemos contextualizar a investigação que proporemos sobre que tipo de texto pode ser considerado culto ou não no século XIX. Traremos respectivamente os textos de Pagotto (1998) e Pagotto (2011) sobre debates a respeito da norma da época e os textos de Coulo et al. (2010), Sílvia & Lopez (2013) e Barbosa (2005), também abordados respectivamente nessa ordem, sobre trabalhos práticos com cartas do século XIX.

2.2.1 A crise sobre uma Norma Linguística do Português Brasileiro por Pagotto (1998, 2011)

Tratando-se do século XIX, temos o problema de esse ser um período sem norma gramatical explícita: havia gramáticas escritas por portugueses preferencialmente e uma constante discussão sobre como deveria ser a norma da língua escrita do português. Para se compreender o que era um redator culto e um não-culto nessa época, é preciso saber qual era a “norma” do período, ou pelo menos, que padrões que os eruditos deveriam e tentavam seguir. Dois textos muito interessantes de Pagotto (1998, 2011) deixam mais claro a dificuldade de se entender a norma que brotava no século XIX no Brasil e a dificuldade que os próprios redatores tinham na época de tentarem seguir esses padrões impostos.

O século XIX é um século de mudanças linguísticas para o português brasileiro principalmente, e também é o momento em que começam a discutir sobre a possível existência dessa língua brasileira, que na época ainda era desconsiderada – no máximo se falava de dialeto brasileiro. Entretanto, quanto mais se analisava a nossa língua, mais era constatado a distância entre as formas usuais no nosso vernáculo e o português exigido na escrita e prescrito nos manuais de gramática. Dessa forma, uma nova norma culta escrita foi

codificada e Pagotto, em seu texto de 1998, enfatiza os discursos que deram sustentação a esta nova codificação, procurando estabelecer uma relação entre ela e a constituição da nacionalidade brasileira.

Nessa fase teria se consolidado a sintaxe na nossa Língua, segundo Pagotto (1998). Para analisar essa mudança, o autor fez a comparação de dois textos do que se costuma ser a expressão mais alta das elites de um país: o texto constitucional. Observando, então, a constituição do Império, de 1824, e a constituição republicana, de 1892, textos que marcaram, respectivamente, o início e o fim do século em questão, constata-se que os dois foram escritos em gramáticas muito diferentes uma da outra. As diferentes marcas gramaticais de cada texto podiam caracterizar o primeiro como escrito em português clássico, e o segundo, escrito no que se tem considerado como a atual norma culta do português.

As características gramaticais encontradas por Pagotto (1998) na constituição do império também são encontradas em outros textos do português clássico, o que atesta o seu caráter de norma culta do período. Por outro lado, muitas dessas formas passaram a ser consideradas como formas populares no Brasil, não pertencente ao prescrito pela norma culta moderna. Isso permite inferir, segundo o autor, que a mudança radical da norma escrita do século XIX representa também a mudança que teve no modo de os sujeitos falantes encararem as formas linguísticas que deveriam usar na escrita.

Certas formas linguísticas caíram em desuso, sendo substituídas por outras da oralidade, inclusive formas estranhas ao nosso atual português, que continua seguindo o seu percurso de mudanças. Língua falada e língua escrita tomam rumos diferentes; e isso não ocorria apenas no Brasil. Como a língua oral de Portugal também estava passando por mudanças, isso resultou em mudanças também na gramática de lá. Ou seja, se por vezes era usado o argumento de que se deveria usar determinada forma linguística porque assim era usado em Portugal, fica evidente a quantidade de opiniões divergentes: tentavam usar como norma algo que também se encontrava no curso de grandes mudanças.

Porém, essa mudança de atitude com relação à língua falada em Portugal teria sido somente um caso de acomodação às formas emergentes? O texto de Pagotto (1998) traz a hipótese de que o Romantismo, por popularizar a literatura e alcançar um maior público, estaria causando uma emergência da burguesia, pois antes a literatura era restrita à nobreza. Já no Brasil, é de grande destaque o trabalho de José de Alencar, não a sua gramática, mas a sua tentativa de dar uma feição própria à gramática no seu texto, a qual conteria elementos

próprios de falantes brasileiros, tanto no léxico quanto na gramática. Com o romantismo sendo bem sucedido em Portugal, em relação a dar novas feições tanto à língua literária como à norma culta, espera-se que aconteça o mesmo no Brasil. Mas porque, apesar dos esforços de Alencar em trazer os reflexos da língua oral ao seu texto, nossa língua terminou o século XIX com uma norma culta ainda mais distante da língua falada e extremamente próxima do português europeu moderno? (PAGOTTO, 1998, p. 56)

Isso trouxe questões contraditórias como a preocupação em romper politicamente com a Europa, mas era necessário construir uma elite à imagem e semelhança da qual a nação se construa – o que não era de fato o que eles desejavam, pois havia o medo de perder essa distinção entre o povo e a elite. A questão da literatura afetou as decisões dessa época, pois ela é o que pode rapidamente materializar uma identidade cultural **que as elites almejam**.

Dessa forma, não seguíamos a norma culta portuguesa por ser origem, ou seja, não era pela questão do purismo, mas porque essa aproximação com a cultura europeia era uma forma de a elite se manter em oposição aos demais segmentos da população. Por todas essas questões, no fim do século XIX, foi quando a Língua Portuguesa no Brasil (doravante PB) esteve mais europeia do que nunca. Isso possibilitava que a elite se distinguisse da minoria: afirmar que o PB era nossa gramática seria nivelar por baixo, e a elite preferia que a norma culta fosse conhecida apenas a partir de rigorosa educação, garantindo, dessa forma, um processo de exclusão.

Foi só no modernismo que houve uma ruptura do binômio norma culta/literatura, o que parece contraditório, mas é explicado na análise feita por Pagotto (1998) do texto de Melo (1972 apud PAGOTTO, 1998). Esse texto foi escrito em uma época em que os debates já estariam esfriando (década de 40); mas que relê José de Alencar, que é o começo dessa polêmica; e que reflete nitidamente o projeto de construção de nação das nossas elites, na sua ambiguidade paradoxal. Essa contradição, então, ocorre porque nessa época, apesar de José de Alencar trazer formas do português popular ao seu texto, isso ainda não era visto como português brasileiro, era visto como o mesmo português europeu **por ambos compartilharem o mesmo sistema**.

Isso se reforça depois, no momento em que Melo (1972 apud PAGOTTO, 1998) afirma que sim, Alencar, de certa forma, reproduziu a língua do Brasil, mas não por ser outra língua, mas por usar do que ele chama de “estilo brasileiro”. Além disso, Melo afirma que é esse estilo que leva alguns “observadores superficiais ou doutrinadores improvisados em

linguistas” (1972, p. 52 apud PAGOTTO, 1998, p. 60) a falarem de língua brasileira. Tudo isso ganha argumento científico, pois Melo estende a ideia de “fala” de Saussure, dizendo que as diferenças entre o PE e o PB ocorrem pela “somatória dos estilos dos membros de uma nação”, sendo assim havia dois estilos e não duas línguas. Com todos esses argumentos há a tentativa tanto de salvar a gramática empregada por Alencar, como, principalmente, mostrar que o autor sequer pensava em escrever em uma língua brasileira. Melo adota a posição evolucionista, falando de abasileiramento da Língua Portuguesa, chegando ao ponto de preconizar a formação de uma língua diferente no Brasil.

Apesar desse texto ser extensamente analisado por Pagotto (1998), ele critica a posição de Melo (1972 apud PAGOTTO, 1998), pois o mesmo se contradiz elogiando a forma que Alencar usa os pronomes em seu texto, que seria diferente do uso do PE, portanto distante daquele português clássico, mas depois diz que isso não é uma liberdade para os ignorantes (PAGOTTO, 1998, p. 64), mas sim uma liberdade estilística. Isso estaria, então, revelando o ideal purista que havia na época, mas que era negado. Dessa forma, a norma culta brasileira sofreu uma manutenção que atendeu a vontade das elites, que era de construir uma identidade nacional, mas manter as diferenças entre as classes.

Ao contrário do texto aqui recém abordado, seu outro texto, *Rui Barbosa e crise normativa brasileira* (PAGOTTO, 2011), não traz tanto as questões literárias, mas sim a grande polêmica centrada mais no âmbito político que surgiu a respeito das emendas criadas por Rui Barbosa para o Código Civil. Apesar de Rui ser um político e jurista, ele debruçou-se na questão linguística.

Assim como outros de sua época, de acordo com Pagotto (2011), Rui Barbosa fazia parte de uma geração que procurava colocar o Brasil no rumo das nações modernas e criticava o ensino do português nas escolas, propondo reformas no ensino primário e secundário. Apesar de sofrer críticas por causa do custo desses projetos, defendia a necessidade de investir em educação como forma de criar no país as condições para o desenvolvimento nacional, através da implementação da educação. Mesmo com toda essa mudança proposta por Rui, ele sofreu muitas críticas pelo seu conservadorismo, pois, ao mesmo tempo que visava defender a língua contra neologismos e estrangeirismo, era a favor dos arcaísmos, pois “a reestruturação de palavras antigas integraria a dinâmica da língua” (PAGOTTO, 2011, p. 6). Também afirma que vários desses termos emergem em textos contemporâneos ou que aparecem até mesmo na fala de pessoas comuns.

Pagotto (2011) contrapõe essa ideia de Rui Barbosa afirmando que era uma tentativa de fazer parecer natural o que não é. Apesar dessas críticas, o autor do texto argumenta que foi o conhecimento de teorias linguísticas de Rui Barbosa “que permitiu a acomodação da unidade linguística normativa com Portugal, pela rejeição dos brasileirismos gramaticais como corrupção da língua e a adoção dos brasileirismos lexicais como expressão do caráter nacional” (PAGOTTO, 2011, p. 8). Nessa época, era defendido um evolucionismo linguístico com a meta de um aprimoramento na sociedade, e esse evolucionismo seria, então, algo natural, daí viria, então, a contradição de Rui Barbosa sobre ser a favor da reintrodução de arcaísmos, mas contra neologismos, sendo que os dois estariam “insuflando o espírito do nosso tempo nessas formas de outrora” (PAGOTTO, 2011, p. 8).

O texto de Pagotto (2011) fala que as obras literárias nem sempre têm a função de confirmar certa forma linguística, mas que às vezes são usadas para justificar uma certa falha cometida, ou para demonstrar a própria instabilidade do cânone literário. Isso é abordado no texto anterior tratado aqui, de Pagotto (1998), no momento em que Melo defende Alencar.

O autor acredita, então, que provavelmente é nessa época que surge a crença de que não sabemos falar nossa própria língua. No período colonial, era importante o domínio do português para se distinguir de indígenas, mestiços e escravos; além disso, torna-se importante, principalmente na segunda metade do século XIX, não só ser bem escolarizado, mas “saber a própria língua”, o que acreditamos se tratar, provavelmente, do que era considerado ser culto.

Essa discussão aconteceu entre o autor do projeto do Código Civil, o jurista Clóvis Beviláqua, o revisor do texto, o gramático Carneiro Ribeiro, o senador Rui Barbosa, que escreveu um parecer sobre esse projeto, e o jornalista José Veríssimo, que escreveu textos que iam contra muitas críticas de Rui. Nesse contexto surge, então, a questão de ideal de nação moderna, que era defendido por muitos: um povo, um território, uma língua. Deixar alterarem a sua língua era como perder parte do território; porém, só os que têm conhecimento do cânone normativo têm esse poder de perda da língua, ou seja, apenas uma parte do povo, os cultos.

Dentro de toda a discussão sobre colocação pronominal, que era um dos tópicos linguísticos mais discutidos na época, há quem se aproximava mais do português clássico, como Carneiro Ribeiro, mas também quem se aproximava mais do PE, como Rui Barbosa, que dessa forma se distanciava do português clássico e da gramática do português brasileiro.

Rui Barbosa apoiava-se na centralização metalinguística e dessa forma mostra um movimento de refacção da norma onde busca princípios racionais que embasem o uso, e não o contrário.

O texto traz diversas citações das discussões entre todos os participantes dessa discussão, citados anteriormente, que tentavam estabelecer alguma norma gramatical. Havia constantemente a tentativa de explicar o seu próprio erro, pois eles viviam a ler e apontar as falhas cometidas um do outro. Dentre essas explicações surge o termo *brasileirismo*, usado por Carneiro, em relação às ênclises em contexto de próclises obrigatórias.

De acordo com Pagotto (2011), as interpretações feitas para esses casos de ênclise seriam uma espécie de hipercorreção, por causa da apropriação da ênclise do português moderno sem que houvesse uma gramática internalizada para sustentar a distinção entre a ênclise possível e aquela não permitida gramaticalmente por certos contextos estruturais. Nesse momento o autor supõe que decorrente ao processo de incorporação de uma nova norma gramatical, inspirada pela norma do português europeu moderno, na primeira metade do século XIX novas construções foram incorporadas como um processo gradual de ruptura com as formas do português clássico.

Nesse momento haveria uma variação maior entre os escritores e indecisões sobre a incorporação do português brasileiro. Já na segunda metade do século XIX, aconteceria uma tomada de consciência desses escritores e de escribas em geral por conta da codificação que os gramáticos brasileiros começaram a fazer. Essa fase explicaria os casos de Carneiro Ribeiro e Rui Barbosa, pois seria um momento tão agudo que textos de um mesmo escritor apresentariam comportamentos normativos diferentes. É justamente nesse período que ocorre acusações sobre contradições, feitas de um escritor para o outro.

Em sua *Réplica*, Rui Barbosa defende-se de Carneiro, dizendo que havia se esquecido momentaneamente de sua regra, não que estava se contradizendo (PAGOTTO, 2011). Isso ocorre, pois, segundo Pagotto (2011) o descuido se coloca no plano do além racional, ou seja, aquele instante em que a razão se distrai e assim, acaba-se cometendo um erro. Já uma contradição indicaria uma falha estrutural, constitutiva, que não faz parte da construção racional que se deseja erigir. Dessa forma, o texto reforça, nos últimos anos do século XIX, a percepção, e principalmente, o exercício de correção normativa passaram a se submeter fortemente à elaboração gramatical metalinguística. Essa instabilidade do cânone durante o século XIX dá espaço para que a codificação gramatical, calcada em princípios científicos,

dite a ordem da língua, salvando, desse modo, os falantes letrados eruditos da terrível incerteza.

Além disso, Pagotto (2011) acredita que, de uma hora pra outra, a elite letrada brasileira teria descoberto que o conhecimento da língua era algo ainda por se fazer. Segundo a hipótese utilizada por ele, não bastava mais o conhecimento dos usos, pela erudição no contato com a leitura de grandes clássicos, o que era muito defendido na época, até mesmo por Rui Barbosa.

O autor (PAGOTTO, 2011) explica que a crise que surgiu em nossa língua pedia pela ação disciplinadora da ciência, pela aplicação metódica de princípios, que mesmo existindo pessoas contra, foi esse o discurso que por fim foi predominado. Mas que, apesar disso, demorou muito até os gramáticos alcançarem um resultado satisfatório em relação às regras de colocação pronominal por conta dos contra-exemplos muito frequentes que apareciam nas obras dos melhores escritores – e a polêmica entre Ribeiro e Rui é apenas um exemplo disso.

Como comentado antes, houve essa quebra entre a norma e os escritos clássicos, algo reconhecido por Carneiro Ribeiro em sua *Tréplica*. Não interessava mais que escritor usasse determinada forma, mas de que maneira seria possível aplicar racionalmente o cânone gramatical da época, que naquele momento passou a ter o rigor da ciência, ser coerente de maneira taxionômica e, desse modo, o ser coerente foi colocado em uso por quem desejava primar pela boa língua. Ou seja, cada vez mais o domínio do próprio idioma estava associado a não só ter conhecimento dos grandes escritores, mas também da estrutura da língua.

A grande polêmica do texto dessa discussão entre Rui Barbosa e os deputados da Assembleia, gira, de fato, em torno dos usos pronominais, mas houve também a discussão sobre a forma adequada do infinitivo flexionado. Nesse assunto também há a oscilação entre o cânone literário e o cânone gramatical, que estaria sendo formado e aplicado com máximo rigor.

Pagotto (2011) diz que esse possivelmente seja um dos exemplos mais notáveis de como a incerteza normativa desse século gerou uma exacerbação da cultura metalinguística. Para reforçar o cânone gramatical, Rui demonstra que, se valendo apenas o uso pelos clássicos, mais de uma regra, referente à questão do emprego do infinitivo e preconizada por outros escritos da época, também deveria ser invalidada.

Dessa forma, Rui Barbosa se aproveita desse ponto para teorizar mais profundamente a respeito do estabelecimento de uma norma que contraste com o uso dos clássicos. Nesse momento, sua elaboração vai ao ponto de colocar a língua em uma perspectiva histórica e evolucionista. Rui Barbosa apresenta uma teoria de mudança para os usos do infinitivo em que o ponto central é a elaboração mais ou menos consciente que os escritores e gramáticos começam a fazer. Visto que isso ocorre dessa forma mais ou menos consciente, surge então uma espécie de competição entre as formas flexionadas e não flexionadas do infinitivo.

Os diversos momentos pelos quais a nossa Língua passou fizeram com que se chegasse, então, no momento em que o uso normatizado refletiria, em termos evolutivos, um estágio mais avançado dela. A plasticidade da Língua seria explorada para corresponder ao melhor emprego das construções a dispor. Nesse caso, a normatização é a intervenção sobre o corpo da língua, algo que naturalmente muda, com o objetivo de se extrair dela o maior equilíbrio possível, e com essa ideia vem o ideal de beleza estética. Com isso fica mais claro a compreensão de porque não desejavam mais seguir o cânone literário:

De nada valem os dados por si sós, se a eles não aplicarmos a elaboração gramatical, no sentido de levar a língua para o caminho correto. Do contrário, se deixarmos tudo livre, se admitirmos como legítimo apenas aquilo que o uso dita – mesmo as práticas da alta literatura -, chegaremos a resultados indesejáveis, pelos padrões estéticos de cada tempo, claro. (PAGOTTO, 2011, p.21)

Os textos de Rui Barbosa pretendem demonstrar para os leitores da época que se devia agir sobre a Língua, tomar uma posição, ou aceitar usos que ofenderiam a boa lógica. Assim mostra-se que a atitude purista – apesar de sempre negada – se desdobra em várias, ora na ideia de trazer novamente à Língua construções antigas, uma forma implícita de novas formas, principalmente nos casos de estrangeirismos, e ora na ideia de domesticar antigos usos e aplicar a eles critérios racionais, próprios do tempo, com o intuito de manter o processo de evolução linguística no controle.

Com todas essas questões tratadas no texto de Pagotto (2011), é possível notar que a elaboração normativa desse século não se deu só por esses aspectos conservadores, mas também por um filtro imposto pela racionalidade e pela moderna ciência, que tinha o intuito de tornar visível o fato de que a Língua não podia mais apoiar-se apenas em um cânone de usos de alta literatura. Deixando isso claro, houve a necessidade de uma intervenção do estudo sistemático, seguindo-se categorias e práticas da moderna ciência da Língua, dando assim, à Língua, um curso seguro para a continuidade de sua evolução – mesmo que no meio disso houvesse todos os deslizos mencionados anteriormente na aplicação da teoria elaborada.

Apesar de toda a discussão em torno do caso de Rui Barbosa ser por si só um bom exemplo de quão complicado é entender a norma do século XIX, o texto traz complementarmente a discussão em torno dos argumentos contra Rui, feito pelos parlamentares. Com isso, veremos que há ainda mais questões a respeito dessa polêmica.

Até esse momento, a relação entre a norma debatida e entre o português brasileiro, que estava ganhando as características gramaticais que contém hoje, foi basicamente de silenciamento. Mas é justamente na contradição, na dificuldade de estabelecer novos padrões e de se situar historicamente frente à língua que o português brasileiro fala, e assim essa relação ganha alguma voz. Essa gramática internalizada, que todos aprendiam em casa, estava justamente perturbando esses julgamentos contraditórios e as tentativas de aplicação de regras (PAGOTTO, 2011, p. 22).

Segundo Pagotto (2011), é o texto da Comissão do Código Civil da Câmara dos Deputados, *Resposta ao Parecer do Senador Ruy Barbosa*, onde se está mais explícito a retomada da ideia de nacionalidade associada à língua. Além disso, é mencionada nesse texto a expressão *dialeto brasileiro*, que, de acordo com a ótica do escolarizado da época, deveria nos trazer independência literária e nos fazer pensar e escrever do nosso modo. Isso diz respeito à retomada do projeto romântico do primeiro reinado que, como identidade, constrói uma nacionalidade baseada na ruptura com a continuidade europeia. As mudanças propostas deveriam, então, segundo os parlamentares, dar uma estabilidade ainda inexistente para a língua, que, para Pagotto (op. cit.), é o mesmo ideal clássico de atingir um ponto ótimo na evolução, em que o dito dialeto brasileiro conseguiria falar igualmente com as demais línguas de civilização.

Os parlamentares, no texto *Resposta ao Parecer do Senador Ruy Barbosa*, abordam duas ideias: a de que é necessário viver o próprio tempo, pois uma nova maneira de ver e pensar surgiu por conta da dinamicidade das inovações, e, dessa forma, colocou o passado em um ponto remoto demais para ser recuperado; e a ideia da conservação das línguas perante as constantes mudanças (PAGOTTO, 2011, p. 24).

Pagotto (2011) analisa, então, um ponto importantíssimo que até então não tinha sido trazido no texto. Na segunda ideia, as pessoas não teriam a ciência e a literatura para conservar a língua, mas, por outro lado, não estariam mais no isolamento e na ignorância. Haveria um estágio intermediário, em que as pessoas se encontrariam numa espécie de limbo da civilização.

Dessa forma entende-se o pensamento da época como um *povo em via de formação*. Essa noção de povo está fortemente ligada à ideia de nação, pretensamente inclusiva, gerada a partir de raça, língua, ciência e literatura. Assim, a identidade como nação estaria ainda indefinida. Com essa visão os parlamentares viam como natural alguns casos de neologismos e a existência de vacilações linguísticas, ao contrário de Rui Barbosa.

Porém, devemos lembrar que toda essa polêmica se dá a respeito do Código Civil, ou seja, se dá em torno da língua escrita – como observado por Pagotto (2011), havia a preocupação de se usar a linguagem de seu tempo, e essa linguagem era a língua escrita. Mesmo que todos da época concordassem que o texto devesse ser bem compreendido por todos, existia a realidade de que não eram todos que sabiam ler – o que se estende até mesmo aos dias de hoje. Dessa forma, o “todos” que se tinha em mente excluía tanto a parte da população que nem sabia ler, como a parte da população que não era nem enquadrada no considerado *dialeto brasileiro*.

Não só nos textos dos parlamentares, mas também nos de Rui Barbosa, é possível perceber como a referência tomada por eles é feita sobre o eixo da erudição. Ou seja, não está em questão o português brasileiro da maior parcela da população, mas sim o português dos pequenos grupos escolarizados.

É possível que essa discriminação com a língua da maioria apareça de forma explícita quando Rui Barbosa usa a expressão *dialeto brasileiro* nos seus textos para desqualificar a linguagem que ele criticava, dizendo que esse dialeto seria algo de quem não sabe sua língua, alguém tem preguiça, ignorância e mau gosto (RUI BARBOSA, 1902 p. 73-74 apud PAGOTTO, 2011, p. 24). Esses que supostamente não soubessem a Língua, não a saberiam por fazerem parte da grande maioria da população que vivia à margem do funcionamento da escrita e dos cânones literários e gramaticais.

No momento em que Rui Barbosa cita diversos escritores, propositalmente brasileiros, de forma positiva, há uma indicação de que brasileiros poderiam sim saber a própria língua, desde que não se entreguem ao sossego e à preguiça. Porém, nessa época, o *povo* ainda estava por se fazer, pois não contava com todos os elementos das civilizações mais adiantadas e não se apresentava racialmente definido. Dessa forma, nos textos de Rui Barbosa, o povo tem, aparentemente, um referente estabilizado. A disjunção com o referente universal acontece por sua relação com o idioma, visto que o Código Civil, deveria ser, no ponto de vista de Rui Barbosa redigido na “forma irrepreensível” do idioma que fala o povo.

Há, assim, uma relação hierárquica em que o idioma falado e o povo se distanciam por causa das normas. Essa relação de inclusão e exclusão fica mais claro na passagem em que Rui Barbosa insiste na desqualificação do texto do Projeto por “brasileiro”, afirmando que o texto só podia estar escrito em *brasileiro* pois em português não estava.

Há também a defesa do vernacular, no sentido clássico, em que Rui Barbosa afirma que não basta apenas ter grande atenção à gramática, mas deve-se também escrever para o povo. Isso traz de novo a dúvida do que queriam dizer com povo nessa época.

Em primeiro lugar, escrever para o povo queria referir-se à pequena parcela da população que era letrada. Os não incluídos nesse conceito são mencionados pelo silêncio, movimento que reaparecerá no emprego da expressão *dialeto brasileiro*. Ademais, quem seria o sujeito desse dialeto? Seriam os eruditos que se distanciam dos bons modelos de linguagem e do gênio da língua, ou seja, existia a intenção de corrigir o rumo das próprias elites eruditas. O conceito de povo, estaria então, ainda mais distante do atual.

A questão do português brasileiro aparentemente não tem espaço algum em toda essa polêmica, sendo, então, uma não língua. Não é como se não a falassem, mas ela é silenciada, falada apenas por aqueles que nem tem o direito de serem incluídos na categoria povo, que ainda estava em formação.

O último aspecto que é importante de se notar nessa polêmica, e que nos faz compreender o quão essa discussão ganhou as proporções que tinha, é a questão nacionalista. Havia toda essa tentativa de colocar a nacionalidade no curso do bom português, o que era entendido como uma totalidade acessível apenas àqueles que se dedicassem a estudar profundamente o seu modo de ser. Um possível desvio para o *dialeto brasileiro* não é a submissão da escrita à gramática do português brasileiro, mas qualquer desvio que colocasse o português usado por aqueles mais escolarizados fora do curso em que se perca a essência do português – que de acordo com Rui Barbosa seria o verdadeiro costume do nosso falar e a língua de nossos pais. É nesse ponto que vemos uma grande diferença entre as críticas de Barbosa em relação ao texto da câmara e as dos parlamentares – que deixam mais explícitos toda a crise normativa da época – em que temos a oposição entre o *dialeto brasileiro* praticado pela elite escolarizada e a Língua Portuguesa universal, pretendida por Rui, atemporal e posta pura em seu gênio.

A grande circulação dos termos **povo** e **nação** dentro dessa polêmica linguística, mostra, explicitamente, a forte relação entre as discussões de cunho normativo e os discursos

em torno da nacionalidade que estavam em confronto na cena política. Todo esse esforço já comentado, de querer colocar o país e a nação nos trilhos da modernidade acontece por conta do momento histórico da época: temos a república recém-implantada e o desafio de acomodação das forças políticas e econômicas após o regime imperial ter sido derrubado. Existe aí, então, uma questão central no trabalho de Pagotto (2011): como desvendar a articulação aparentemente existente entre o ideal de modernização e de civilização, que perpassa o fim do século XIX e chega ao seu auge na instituição do regime republicano; e o conservadorismo linguístico presente nesse mesmo período, que tem como ápice a polêmica do Código Civil?

Nesse ponto surge a dúvida de como Rui Barbosa, tendo proposto reformas no ensino e tendo se mostrado extremamente por dentro das tendências modernas de educação, e até mesmo outros ativistas de sua geração, tendo tomado contato com uma variedade de perspectivas modernas de ciência, teriam sido tão conservadores em matéria de língua. Essa contradição existe no momento em que querem moldar a nossa língua a partir dos mais escolarizados para os menos escolarizados.

Existia essa ideia de seguir tendências científicas modernas, pois era isso que era feito na Europa, mas justamente por isso era necessário reforçar nossa filiação com a origem da nossa língua – o português europeu – e dessa forma tentar moldar nossa língua de acordo com o que se passava lá, e não com o rumo natural que nossa língua estava tomando. Trata-se, então, de um movimento de desvinculação da nacionalidade e desse suposto dialeto brasileiro, tomado não apenas como sujeito às influências do português do Brasil, mas também à mercê de todos os estrangeirismos e neologismos que circulavam entre as camadas mais escolarizadas da população.

Nesse ponto é importante retomar. Esse suposto dialeto brasileiro seria os desvios cometidos pelos já escolarizados, já que a língua falada de fato pela grande maioria da população nem faz parte dessa discussão. A questão nacionalista fica extremamente clara no momento em que a única nacionalidade possível de ser expressada é essa língua depurada. Sendo assim, trair essa língua é como trair a própria nação e defendê-la de estrangeirismos, é, portanto, defender a nossa própria nacionalidade (PAGOTTO, 2011, p. 28).

Mas essa questão traz de novo o aspecto excludente que vimos nessa discussão, pois há a afirmação de que não sabíamos nossa língua – feita pelo jornalista José Veríssimo nos textos da polêmica do século XIX. Existe esse trabalho de desvincular a nacionalidade de uma

pretensa expressão linguística brasileira, colocada no trilho da língua única, única capaz de expressar corretamente essa nacionalidade que ainda estava por se fazer.

Porém já vimos que quem sabe essa língua não são todos, então é necessário aprendê-la e cultivá-la. Apenas dessa forma teríamos uma nação plena. Isso nos leva, então, ao fato de que a nacionalidade era constituída a partir de uma posição de saber, e ressaltado, então, o fato de que essa posição do saber não era ocupada completamente nem mesmo pelos mais eruditos – e quem dirá, então, o restante da população. A crise em torno da identidade nacional fica clara com essa afirmação de Veríssimo, onde se é instaurado a noção de que não se pode ser sem conhecer.

Isso nos leva a outro aspecto importante em toda essa polêmica: a necessidade de se manter uma divisão clara da sociedade entre povo e elite. Para entender essa questão o texto traz o trabalho de Alonso (2001 apud PAGOTTO, 2011), que faz uma análise da crise política no final do Império e da geração de intelectuais e políticos que dela participaram. O trabalho de Alonso é citado para dialogar com o movimento arcaizante de Barbosa e outros autores da época, que preveem uma refundação da nação sobre a qual não se perca o controle. A intenção dos autores do Código Civil era de reafirmar a filiação europeia, pois o trabalho de Alonso (2001 apud PAGOTTO, 2011) trata justamente da questão da nação escravista, tecnologicamente atrasada, fundada na dinâmica do grande proprietário de terra.

Alonso (2001 apud PAGOTTO, 2011) traz um estudo detalhado da crise política que culmina na queda do Império, tendo como hipótese central que se tratou de um movimento de diversos grupos de excluídos do sistema político da época. Esses grupos agiam a partir de elaborações teóricas modernas e progressistas, que estavam repletas, nesse caso, de uma práxis que visava a uma remodelação do sistema de poder e a um novo projeto de nação.

Dentre as interpretações desse período, chama-se a atenção para a superficialidade com que as novas teorias são tratadas pela então inteligência brasileira, que teria reduzido o conhecimento estrangeiro a um mero verniz de erudição com que se procurava compreender o País e a Nação, porém, sem entrar de forma profunda no mérito do seu entendimento. Entretanto, Alonso (2001 apud PAGOTTO, 2011) rejeita essa interpretação, e propõe que a geração contestadora de 1870 deve ser entendida como um grande grupo da elite escolarizada que estava, de algum modo, fora das formas de inserção política.

Os grupos contestadores dessa época se diferenciavam ou se aproximavam uns dos outros em função do tipo de proposta para a estruturação política do país, assim como em

função da análise das causas do atraso do país e a visão que tinham de nação. O que os aproximava era a experiência da marginalização política e, como plataforma, a causa abolicionista, mas nem sempre pelas mesmas razões.

E é a partir disso que compreendemos a questão trazida no parágrafo anterior. Todos esses grupos contestavam o projeto saquarema de nação, que caracteriza o segundo reinado (o indianismo romântico, liberalismo estamental e o catolicismo hierárquico). Eles buscavam reformar o estado em bases modernas, laicizando-o e, ao mesmo tempo, estabelecer um novo projeto de nacionalidade, de natureza mais inclusiva. Aprofundar a nacionalidade implica investir contra o ideal de nação calcado na especificidade da nacionalidade brasileira reivindicado pelos ideólogos do império, que aparece manifesta na associação entre o branco europeu e o índio. Ou seja, era uma questão de abandonar o ideário indianista de nação e buscar aproximar o Brasil do mundo europeu. Então, existe, dessa forma, a procura em minar o projeto político saquarema, a cuja limitação se debita o atraso do país que demandava por reformas. O moderno desejado buscará na cientificidade da época as bases intelectuais para pensar o Brasil de maneira prática, buscando soluções para o impasse colocado, impasse esse que Pagotto, referente ao texto de Alonso (2002 apud PAGOTTO, 2011), explica:

Na interpretação da autora e de outros, a crise que derrubou o império foi uma crise intra-elites: no modelo adotado no período da Regência, a representação política e a inserção nas estruturas de poder estava fortemente centrada na posse da propriedade de terras. A partir da modernização do país, uma geração de despossuídos com formação intelectual, mas sem o lastro da grande propriedade se sente alijada das relações de poder e reclama reformas. (PAGOTTO, 2011, p. 29)

Dessa forma entendemos, então, toda a problemática comentada no texto a respeito do uso das palavras **povo**, **nação** e **dialeto brasileiro**. Só pertencia aos dois primeiros conceitos, aqueles com capital de conhecimento o suficiente, ou seja, aqueles que se submetiam a essa linha evolutiva colocada pelo conhecimento da tradição erudita, e não, necessariamente, aqueles com herança nobre de grandes famílias. Bastava, dessa forma, essa submissão, para integrar da maneira mais profunda a nacionalidade; praticar a língua numa determinada direção, que era ditada pela ancestralidade.

O dito dialeto brasileiro, termo muito usado durante a polêmica do código civil para desqualificar o texto, não chegava nem perto da língua real, falada na época. Ou seja, existia na época toda uma gama da população (a sua maioria) que não fazia parte dessa discussão, por não fazerem parte nem da língua considerada errada, e, portanto, eram excluídos da própria nacionalidade.

Através dessa grande polêmica minuciosamente abordada no texto de Pagotto (2011) é fácil perceber por que o autor se refere ao período como uma “crise normativa”. A partir de um documento, que deveria ser revisado do ponto de vista político, e que foi revisado, na verdade, praticamente apenas do ponto de vista gramatical, surgiu uma discussão em que todos os intelectuais da época tinham algo a acrescentar – e que acreditavam que era o ponto de vista deles o certo, e não o do outro, muito menos o dos falantes em geral. Com isso vemos como a língua sofre pressões não apenas sociais, mas também políticas:

A noção de certo e errado que se imprime a esta ou aquela construção, seria, assim, fruto de uma operação simbólica em que o político, o ideológico determinam as significações sociais que passam a operar no jogo sociolinguístico propriamente dito (PAGOTTO, 2011, p. 31).

A construção social de que “nós não sabemos nossa língua” perdura desde essa época, através de uma polêmica derivada de um purismo exacerbado e um zelo excessivo no exercício do ideal clássico, derivados dessa crise normativa. O autor refere a crise normativa como “a tomada de consciência da mudança nos padrões normativos” (PAGOTTO, 2011, p. 31), que ocorre justamente entre os anos 1850 e o episódio da polêmica do código civil, um momento em que o que se tinha em mente como padrão deveria mudar. Foi apenas após essa polêmica que o “dialeto brasileiro”, falado pelas elites que buscavam estabilidade, passa a se referir a língua de fato usada pelos brasileiros excluídos dessa crise.

2.2.2 Outros estudos sobre o perfil social do redator de sincronias passadas

A sessão anterior deixa evidente as dificuldades que um pesquisador irá enfrentar ao tentar fazer uma pesquisa Sociolinguística em séculos passados, como no nosso caso, o século XIX. Como será melhor explicado pelo texto *Reflexões metodológicas para a análise sociocultural de redatores em corpora históricos* (COULO et al., 2010) a importância de se saber qual era o padrão de escrita usado em uma determinada época vem da necessidade que a sociolinguística tem de fazer seu estudo a partir do vernáculo. Uma vez que o vernáculo surge em situações reais de fala já se torna mais difícil obtê-lo em fontes escritas, que em geral toma-se mais cuidado na forma de se expressar e se usa a fala mais formal. Além disso, a partir do momento que não temos uma norma bem definida no século XIX, como foi extensamente exemplificado anteriormente, fica claro como existe uma enorme barreira para o pesquisador.

Diversos autores, como o já mencionado Coulo et al. (2010) e Barbosa (2005), concordam que o modo mais aproximado de conseguir o vernáculo, seria, então, através de cartas pessoais, onde os redatores não utilizariam uma linguagem tão formal. A seguir explicaremos as dificuldades encontradas por esses e outros autores, como o fato de pessoas com alto grau de escolaridade (ou, como usado para séculos passados, alto grau de cultura) tenderem a ter um cuidado muito grande mesmo em textos mais informais; e com isso surge a dificuldade de de que modo diferenciar o formal e informal do culto não-culto.

O texto de Coulo et al. (2010) aborda algumas questões metodológicas complexas, como a ideia de o indivíduo refletir a sociedade em que vive, questionando então os conceitos de indivíduo e sociedade. O autor também releva muito a questão de que, além de ser necessário ter todo um perfil sociocultural do redator, é preciso saber o contexto em que determinado texto foi escrito, pois essa informação, juntamente com os dados da trajetória de vida dos redatores e o mapeamento e descrição das redes escritas possibilitaria uma conceituação alterna de perfil sociocultural. Para isso, Coulo (op. cit.) ressalta a importância de se trabalhar com o gênero epistolar¹, visto que se trata de uma circunstância espontânea de comunicação verbal.

Coulo et al. (2010) afirma que a estrutura fixa desse gênero textual, que começa com local, data e saudação inicial, permite a identificação do perfil sociocultural do remetente na sua relação com o destinatário, porém, ainda assim não reproduzem dados de fala. Além disso, apesar da carta, com seu caráter mais íntimo ou espontâneo, transmitir a inovação e mudança linguísticas, ela também conserva fórmulas fixas de enunciados. Sendo assim, o pesquisador tem mais um desafio: o de observar até que ponto a identificação de um dado mais conservador ou inovador daquele material evidencia de fato uma etapa de um processo de mudança sistêmica da língua ou se trata apenas de uma peculiaridade daquele tipo de texto, no caso, o gênero epistolar.

Marcotulio (2008 apud COULO et al., 2010), ao analisar as cartas do Marquês do Lavradio, português e vice-rei do Brasil, escritas no Rio de Janeiro, no terceiro quartel do século XVIII, trocadas com diferentes destinatários, inclusive familiares, notou que as relações de parentesco indicadas pelos rótulos usados por remetente e destinatário muitas vezes não correspondiam às informações conhecidas daquela família. O cargo ocupado por

¹ No texto, a palavra epistolar surge simplesmente como um sinônimo de cartas. Em dicionários encontraremos diversos outros significados, até mesmo religiosos. Em geral epístola é usado como carta ou se refere a algo que tenha relação a cartas.

Lavrado deveria ser de subordinado somente à Coroa, mas deveria ter poder sob os governadores e capitães-generais das demais capitanias do Brasil, porém, para garantir que o poder continuasse do lado europeu não era deixado claro os reais limites de atuação do vice-rei, fazendo dele então, mais uma figura representativa do que de fato vice-rei.

Dessa forma, Marcotulio (2008 apud COULO et al., 2010) decidiu estudar os pronomes de tratamento usados nas cartas de Lavradio, por acreditar que elas forneciam indícios de como as relações entre formas linguísticas e papéis sociais se construíram ao longo dos tempos. Foi constatado, então, que eventualmente o destinatário ou remetente subvertia/ampliava os limites dos laços familiares tais como são conhecidos hoje, como por exemplo chamar os genros de “filhos” ou sogra de “mãe” ou até mesmo um amigo muito próximo de “irmão” - transformando dessa forma a esfera familiar em algo além de laços sanguíneos ou matrimoniais.

Outro trabalho que demonstra preocupação com a identificação de categorias sociais é o de Rumeu (2008 apud COULO et al., 2010), que fornece, entre outros aspectos, passos metodológicos para a identificação dos perfis socioculturais dos redatores brasileiros das epístolas familiares trocadas entre os entes da família Pedreira Ferraz-Magalhães na realidade sócio-histórica entre os fins do século XIX até metade do século XX. Através de buscas em arquivos públicos, o autor conseguiu descobrir dados de cada pessoa da família, como grau de escolaridade e também saber o grau de parentesco entre os redatores. Esse estudo possibilitou resgatar a progressão da mudança linguística (a implementação de *Você* no sistema pronominal do PB) através da análise do comportamento linguístico do indivíduo graças a observação da produção escrita em distintas faixas etárias da vida do redator.

Por fim, Rumeu (op. cit.) cita um terceiro trabalho, dessa vez com folhetins e não cartas pessoais. A importância de se trabalhar com esse tipo de texto vem do fato desse material ter servido de modelo de erudição para o século XIX. Como Lima (2010, COULO et al., 2010) procurou ratificar, os impressos em periódicos faziam parte do cotidiano do carioca, que além de ter a função de entreter e informar, formava os redatores e leitores da época. Porém o desafio dessa pesquisa consistiu em conseguir identificar quem era o autor do texto, a sua escolaridade e origem, na ausência de original preservado e de informações sobre o alcance desses textos. Todas essas informações tidas para Coulo et al. (2010) como condições essenciais para se obter qualquer afirmação sobre a norma pertencente àquele grupo. Nesse contexto Lima escolheu um redator, França Junior, o qual se tinha mais informações

socioculturais e que, devido a popularidade de seus textos, provavelmente se enquadrava dentro do padrão normativo social da época.

Esses dois trabalhos e o com os folhetins dependem, entre outras coisas, de saber sobre a escolaridade do redator. No caso do estudo com os folhetins, quando não encontravam nenhuma outra informação sobre o autor, foram obrigados a descartar aqueles textos. Tudo isso mostra a importância de uma metodologia para se descobrir o perfil sociocultural dos redatores, fato que Coulo et al. (2010) também concorda no momento em que cita o trabalho de Lima (2010 apud COULO et al., 2010) como colaborador da metodologia de Barbosa. Esse trabalho, por levantar mais um informante de grau de cultura elevado, França Junior, para compor o parâmetro de erudição e testar em seus folhetins a metodologia de Barbosa (2005), está corroborando com a ideia desse trabalho que será melhor explicado agora.

O trabalho de Barbosa (2005) exemplifica de forma prática os problemas de se trabalhar com textos de sincronias passadas. Se trata das (como nomeado por ele) *cartas dos avós*, redatores de terceira idade, identificados como de classe alta, reconhecidamente cultos para a época, que trocavam cartas dentro do círculo familiar e são divididos pelo fator sexo. O autor explica, que mesmo tendo um *corpus* com muitos pontos positivos onde se sabe quem eram os redatores, naturalidade deles, de que tradição discursiva se tratava, possibilidade de contrapor homens e mulheres, etc. ainda existe o problema em torno do significativo *versus* representativo. De fato, o *corpus* é de extrema significância, mas o autor levanta a dúvida de ele realmente representar os falantes de terceira idade da época.

Hoje em dia conseguimos facilmente identificar um discurso formal de um discurso informal, pois vivemos na mesma época e compartilhamos dos valores de nossas normas objetiva e subjetiva, mas podemos erroneamente julgar o informal de épocas passadas como formal. Tanto a formalidade e a norma culta como a informalidade e a norma não culta parecem caminhar bem próximas no plano da escrita cotidiana. Sendo assim, é preciso se saber o que se compreendia ou pelo menos se praticava como informal naquela determinada época, ou seja, precisamos saber qual era o modelo de cultura escrita seguido na época para ser possível identificar quando ele se afastava desse modelo ao ponto de se tratar de um discurso informal mais próximo de uma escrita cotidiana - dessa forma se aproximando mais da oralidade e, podemos supor, com usos linguísticos mais próximos da norma geral vernácula.

Esse é um dos problemas enfrentados na pesquisa de linguística histórica, onde surge a necessidade de operar com critérios mais objetivos, época a época, o grau de formalidade do redator, tanto em relação ao distanciamento entre emissor e receptor como ao grau de cultura dos materiais que os pesquisadores têm acesso. Por conta disso que Barbosa propõe um controle mais seguro para se obter essas informações socioculturais do que meramente nossa intuição, construindo então, um provável parâmetro de norma culta objetiva para o século XIX no Brasil - ou seja, não a ensinada na escola, mas sim aquela efetivamente praticada pela sociedade. Ele usa como base as cartas dos avós Ottoni, redatores cultos da época, e também se questiona se os aspectos morfossintáticos corresponderiam aos aspectos gráficos - e buscando responder isso analisa as formas sintéticas e analíticas do gerúndio.

Sendo assim, o ideal seria trabalhar com texto que se tem algum conhecimento do redator, para conseguir, de forma aproximada, saber se naquele caso ele estava usando uma escrita mais ou menos formal. Como em geral não se tem o perfil sociocultural dos redatores, o primeiro passo, é na verdade, descobrir o seu grau de cultura - como dito antes, não o grau escolar dele, pois essa seria a norma culta escrita subjetiva, encontrada também nas gramáticas, mas sim a norma culta escrita objetiva. Formalidade pode ser entendida tanto como adequação a um modelo, convenção, como a uma atitude de distanciamento entre emissor e receptor. Logo, deve-se manter em mente que mesmo um redator culto, escrevendo um bilhete ou algo que não requer tanta formalidade, deixará alguma marca daquilo que é considerado padrão culto para sua época. Do contrário de um redator sem conhecimentos do padrão da época, que mesmo na tentativa de escrever um texto formal dentro dos padrões, deixará evidente, pelos seus *desvios*, que não tem contato com os gêneros discursivos que servem de modelo para os eruditos de sua comunidade. Ou seja, o domínio daquilo que se entende como norma culta interfere na real formalidade de um determinado material escrito.

Fatores necessários para, teoricamente, se saber o quão culto um redator é, são citados no trabalho de Barbosa (2005): vislumbrar os valores de erudição da época que difiram dos da nossa, reconhecer os modelos textuais de padrão culto objetivamente à disposição, e, determinar um referencial indubitavelmente identificado como marca de prestígio na cultura escrita da época e lugar em questão. O autor toma como textos modelos os jornais da época, e como marca objetiva, para determinar um padrão quantitativo, o uso de palavras etimologizadas. Por conta dessa busca pela marca de prestígio culto, muitos redatores não tão cultos acabavam por fazer o uso de forma errônea, ou seja, grafavam palavras que não eram de origem latina (ou grega) - isso também é levado em consideração na metodologia do autor.

O texto confirma a hipótese dos textos padrões da época no momento em que contrasta com um redator de perfil sociocultural conhecido: o avô Ottoni, professor, engenheiro e senador, teve um uso de 2,81% de latinizações, e taxa de acerto de 92,15%; em contraste com 3,69% de uso e 90,35% de acerto nos jornais. Através dessa metodologia o autor propõe três graus de cultura para a escrita oitocentista: *erudito*, *informante culto* e *informante de cultura mediana*. Abaixo e fora desse eixo de cultura estariam os *semicultos* e as chamadas *mãos inábeis*². Nessa classificação o avô Ottoni seria considerado erudito, enquanto a avó Ottoni, com um uso mais baixo, de 1,29% entraria na faixa dos informantes de cultura mediana - apesar de acertar todos latinismos usados. O autor explica que a taxa de acerto dela não pode ser levada em consideração devido aos outros erros cometidos por ela.

Para confirmar uma dúvida lançada no início de seu texto, o autor traz, por fim, os dados a respeito da análise morfossintática: uso de formas analíticas e sintéticas do gerúndio. O levantamento é feito novamente contrastando o uso dos avós com o dos textos de jornais: 90,31% de formas sintéticas e 9,68% de analíticas em jornais contra 64,51% de formas sintéticas e 35,48% de analíticas usadas pelos avós. Porém, analisando os avós separadamente, vemos que o avô (70,37% de uso da forma sintética e 29,62% da analítica) se aproxima mais dos padrões dos jornais do que a avó (25% de uso da forma sintética e 75% da analítica). Ou seja, esses resultados reforçam, agora em dimensão gramatical, a distinção já feita em nível grafológico.

Além disso, os resultados evidenciam que, segundo o recorte aplicado, as cartas dos avós revelam-se mais próximas do padrão oral do que os periódicos impressos, uma vez que formas sintéticas são tidas como mais usuais na escrita e as analíticas mais frequentes na fala. É importante ressaltar, então, que o redator, que tem mais chances de apresentar de fato, uma escrita cotidiana mais aberta aos possíveis usos já generalizados na sociedade da época, é, portanto, a avó Ottoni. Isso mostra a tamanha significância desse trabalho, que consegue propor duas metodologias para se classificar não só o grau de cultura do redator (através do uso de etimologizações) como também o grau de formalidade de um texto (através do uso de formas sintéticas contrastadas com formas analíticas) – apesar dessa última metodologia não ter sido extensamente testada e ter de ser analisada sempre levando em consideração o grau de cultura do redator, para não cair na problemática culto/não-culto *versus* formal/informal e julgar erroneamente o missivista.

² Cf. MARQUILHAS, 2000

Outro trabalho interessante a respeito de metodologias para determinação de perfil sociocultural de redatores antigos é o de Sílvia & Lopes (2013). O *corpus* utilizado é composto por cartas de um casal não ilustre de 1930, Rio de Janeiro. Por se tratar de um casal não ilustre não foi possível recuperar informações sobre eles - por isso a necessidade de uma metodologia para definir isso. As autoras utilizam o termo *grau de letramento*, que seria determinado pelo desvio grafemático; etimologizações; abreviações e segmentação/junção de palavras. Ou seja, é feita a análise do vocabulário e da estrutura sintática, ortografia, pontuação e segurança do remetente das cartas. Essas informações são processadas através do programa *E-dictor*.

Na primeira tabela dos resultados as autoras separam por menor domínio de escrita os critérios de segmentação/junção e modernização (grafia), e por maior domínio de escrita modernização (etimologização) e expansão (abreviaturas). Em todos os fatores a redatora feminina apresenta menor conhecimento da escrita (maior porcentagem que o redator masculino nos critérios de menor domínio de escrita e porcentagem inferior nos critérios de maior domínio de escrita. É feita uma análise detalhada de todos os critérios analisados pelo *E-Dictor*: segmentação/junção, desvios grafemáticos, monotongação e ditongação (alternância entre [e] e [i] e entre [o] e [u]), apagamento da nasalidade, síncope de vogais e consoantes, etimologização e expansão das abreviaturas.

O texto conclui que apesar de haver uma diferença entre o grau de habilidade dos missivistas no que concerne às normas de escrita da época, não se pode afirmar que eles ocupam posições extremas, pois ambos apresentam evidências de transposição de aspectos fonético-fonológico nas cartas. A partir da análise foi identificadas diferenças quanto ao grau de letramento de cada redator e, a partir disso foi proposto uma escala de cinco níveis a partir da dicotomia [-] letrado para [+] letrado, ficando o redator em segundo nível mais alto e a redatora em quarto.

A tabela tem breve descrição de cada nível, o nível ocupado pelo redator tem como características “Presença de desvios grafemáticos em baixa ocorrência motivados pela oralidade a palavras pouco usuais. Presença de falsas etimologizações em baixa frequência; Ausência de segmentação/junção de sílabas ou palavras.”, e o nível da redatora é caracterizado por “Repetição e pouca variabilidade de etimologizações e presença de pseudoetimologizações; Segmentação/junção de palavras de maneira sistemática tanto em formas presas quanto em fronteiras silábicas. Desvios motivados pela oralidade, apresentando muitas oscilações na grafia de uma mesma palavra;”. Dessa forma, constatou-se que o redator

masculino detinha um nível de conhecimento relativo a textos escritos e um maior contato com modelos de escrita que sua noite, visto que ela apresentava mais ocorrência de desvio grafemáticos e segmentação/junção de palavras assim como uso indevido de vocábulos desnasalizados, monotongados e ditongados, que trazem traços de oralidade para a escrita.

As autoras fazem um comparativo, ao final da conclusão, da noiva com a metodologia de Barbosa (2005), colocando-a na categoria de *culta-mediana*. Acredito que isso esteja equivocado por dois fatores: primeiro que se trata de séculos distintos, como visto anteriormente na polêmica do código civil, ao fim do século XIX, início do século XX houve estabelecimento de certos padrões ainda não bem estáveis; segundo que essa redatora cometia muito mais erros do que a avó Ottoni, principalmente referentes a resquícios da oralidade, mostrando que seu conhecimento da norma culta escrita era de fato inferior do da avó Ottoni. Apesar de se tratar de fato de mais uma importante pesquisa em busca de metodologias para se determinar o perfil sociocultural de redatores desconhecidos, as autoras utilizam uma metodologia praticamente qualitativa, extremamente minuciosa e, que eu acredito, ser de difícil aplicação em *corporas* de grande volume.

Ao se trabalhar com textos históricos fica evidente a necessidade de se saber, como já dito no início da seção, a norma daquela época. No nosso caso, nos deparamos com a falta de uma norma bem definida do século XIX. É importante retomar, que essa não definição de uma norma não é um problema apenas para nós pesquisadores do século XXI, mas era uma dificuldade também dos redatores daquela época, que não sabiam exatamente que padrão seguir.

Os trabalhos práticos de Barbosa (2005) e Sílvia e Lopes (2013) são ótimos exemplos dos obstáculos encontrados por pesquisadores que decidem por trabalhar com textos antigos. Barbosa (2005) mostrou justamente a busca pelo padrão seguido na época, através da hipótese e confirmação de os periódicos dos jornais representarem o padrão de escrita seguido na época. Além disso, por dispor de um *corpus* muito importante, onde se tinha uma grande informação sobre os redatores, também colaborou com a hipótese de se descobrir um método objetivo para saber os perfis sociológicos de *corporas* de redatores desconhecidos.

Sílvia e Lopes (2013), também com esse objetivo de desenvolver uma metodologia para se ter o conhecimento do perfil dos redatores, se diferenciam por apresentarem uma proposta não tão objetiva. Esses dois trabalhos, além de mostrarem a importância de se saber

o perfil dos redatores, e se desenvolver um método para isso, nos mostra como ainda há o que ser pesquisado nessa área, uma vez que se tem mais dúvidas do que certezas.

3 DIRETRIZES DA PESQUISA

Comentada toda parte teórica, partiremos para nossa parte prática. Além de expormos melhor nossos objetivos e hipóteses, demonstrando o porquê da escolha desse tema e sua importância, explicaremos a metodologia utilizada na tentativa de alcançarmos isso. Como o material analisado consiste em dois *corpora* distintos, daremos informações mais pontuais sobre ambos.

3.1 OBJETIVOS E HIPÓTESES

Dentre os diversos desafios da sociolinguística histórica em geral, como o de se identificar social e culturalmente um indivíduo de certa comunidade linguística em um dado estágio de língua, por sua linguagem espontânea e natural, esse trabalho propõe a construção de uma metodologia preliminar de busca de textos que representem o vernáculo.

Visto que mesmo hoje em dia, com muito mais recursos que antigamente, há barreiras em estudá-lo, ressaltamos o fato de que há ainda mais problemas em se estudar o vernáculo do passado, tendo em vista que é desconhecido. Apesar de falarmos tanto de dificuldades, é necessário reconhecer que há sim uma quantidade significativa de material em museus e arquivos públicos para estudo. O problema da utilização desse material para estudos linguísticos é justamente a dificuldade de se obter um texto que transpareça uma linguagem menos monitorada, uma linguagem mais espontânea e natural.

Isso ocorre basicamente por dois motivos: o primeiro é que a maior parte dos materiais encontrados em museus são cartas trocadas entre pessoas de alto nível social (entre duas pessoas pertencentes a elite), ou seja, será uma linguagem bem monitorada, não espontânea; o segundo é a quantidade de cartas que encontramos sem informações socioculturais dos redatores (fator essencial para a pesquisa em sociolinguística). Em geral, documentos que despertam mais interesse por parte de pesquisadores diversos (principalmente da área de história e sociologia) são justamente cartas de grandes personalidades históricas - portanto pessoas que tiveram acesso a estudos.

Quando não se trata desses casos, são cartas de pessoas desconhecidas, ou seja, pouco se sabe se aquela carta tem o propósito de ser formal, mas nos parece informal por ser uma pessoa menos familiarizada com a escrita ou se é uma carta informal, que parece formal por

se tratar de uma pessoa muito erudita. Sendo assim, os dois casos, sem maiores investigações, podem nos trazer uma falsa ideia do que era vernacular naquela época.

Barbosa (2005) explica a confusão que se pode fazer entre *formal x informal* e *culto x não culto* em textos antigos. Em um texto contemporâneo, conseguimos dizer empiricamente se ele é um texto formal ou informal, mas isso ocorre pois vivemos na mesma época desse texto e temos conhecimento dos valores de nossas normas objetivas e subjetivas. Porém, nos textos de antigamente, “podemos ter um informante culto em atitude informal (ou outro, de menor erudição, sendo formal quando lhe aprouver)[...]” e não reconhecer isso com a mesma facilidade, pois “[...] formalidade e normal culta, tal como informalidade e norma não culta parecem caminhar bem próximas no plano da escrita cotidiana.” (BARBOSA, 2008, p. 27).

Além disso, por não sabermos os padrões de cultura escrita desse período e o quanto a dicotomia entre a fala e a escrita foi aumentando, “[...] não sabemos o que era ser culto no passado, tampouco se o que sentimos como formal, fosse, para eles, informalidade.” (BARBOSA, 2005, p. 28). Portanto, sem um método objetivo essa distinção se tornaria muito arbitrária.

Ou seja, temos a dificuldade não só de encontrar cartas de cunho mais pessoal e íntimo, visto que são personalidades históricas os principais redatores, mas também e normalmente há uma preocupação em cima desse material, tendo em vista que as cartas são trocadas entre personalidades históricas e políticas também de grande importância (resultando, então, numa conversa formal). Há também o fato de que, mesmo em conversas com entes queridos, o texto apresentar-se com certa formalidade nos dias atuais, por causa do alto grau de cultura que possuem, além de hierarquias sociais diferentes na época (filho deveria ser muito formal com seu pai, por exemplo, por questão de respeito).

Uma ideia para contornar esse problema de identificação do que seja um texto formal ou não no século XIX, seria a de se buscar cartas de pessoas mais comuns, sem tanta escolaridade, trocadas entre os familiares, que tratassem de assuntos mais banais - tudo para garantir o exame de um possível vernáculo. Porém, isso nos leva ao segundo motivo que impossibilita o estudo sociolinguístico de uma grande quantidade de cartas.

Sendo essas pessoas mais comuns, as informações que se tem delas são menores. Não se sabe idade, grau de cultura (que é a informação que nos interessa nesse trabalho) e por vezes nem o gênero do redator, visto que um texto (nos casos de possível desconhecimento, de fato, da norma escrita) poderia não ser escrito pelo próprio signatário, autor no corpo da

carta, o que dificulta saber os papéis sociais entre esses redatores, dado importante para estudos como o de Marcotulio (2008 apud COULO et al., 2010), mencionado na seção 2.2.2.

Desse modo, temos acervos de importância histórica elevada, mas que pecam no quesito de se poder trabalhar com o vernáculo: a língua em situações reais de uso. Os documentos antigos podem ter um grande conteúdo possivelmente informal, mas que, por não se ter informações socioculturais concretas sobre os redatores, não é possível ser usado por não se saber o grau de cultura dos redatores.

Esse trabalho tem, portanto, o objetivo geral de se propor uma metodologia para se tornar mais utilizável parte dos *corpora* disponíveis em Arquivos Públicos, para pesquisa linguística histórica. Além desses *corpora* terem mais chances de apresentarem vernáculos eles também podem apresentar traços de oralidade, como por exemplo no caso das cartas dos avós Ottoni³. Como reconhecido por Lopes (2005) “[...] nas cartas de Bárbara identifica-se a presença de traços de oralidade”, fato que faz com que a autora considere a avó Ottoni como “atriz principal” desse *corpus* (p. 18). Porém, como já frisado, não encontramos tantos casos como o de Bárbara.

Buscamos esse objetivo, inspirados em um trabalho realizado por Barbosa (2005), levando em consideração duas de suas hipóteses. Primeiro a de que jornais de época sejam textos-modelos. A partir dessa premissa, Barbosa (2005), afirma que nesses textos modelos deveria ser possível notar alguma marca padrão. Essa seria a sua segunda hipótese.

Apesar de o autor ter suas duas hipóteses como comprovadas, mas mantê-las “sob a normal suspeita que paira sobre as investigações científicas em fase inicial” (p. 41), partimos do pressuposto de que sim, os jornais marcavam a norma culta escrita da época - pois o teste foi feito, na sua pesquisa, com um número muito significativo de 42.269 palavras.

Já essa segunda hipótese é a que de fato nos debruçamos a aperfeiçoar no presente trabalho. Essa marca padrão foi identificada no uso intensivo de etimologizações, sendo esse, então, considerado o marcador de grau de cultura nos textos do século XIX. Ao contrário da primeira hipótese que não discutiremos nesse trabalho, acreditamos que a segunda deve ser mais explorada, visto que foi examinada por Barbosa (2005) com um número pequeno de redatores (apenas o avô e a avó Ottoni) e, portanto, possivelmente não tenha sido

³ Os avós Ottoni foram o casal Christiano Benedicto Ottoni e Barbara Balbina de Araújo Maia. Ele foi militar, lecionou matemática, tornando-se depois catedrático, formado em Engenharia, também atuou como político e escritor. Sobre ela há poucas informações, supõe que foi alfabetizada pelos ditos “professores domésticos” e sabe-se que seus pais possuíam “fortuna média”.

suficientemente cumprido tão bem o critério objetividade - uma vez que não teve nenhuma fórmula classificatória ao final do texto.

Uma vez que o objetivo é a hipótese de se testar o critério das etimologizações, ela deve ser melhor explicada. A complexidade da trajetória da Língua Portuguesa até a língua que conhecemos hoje e as polêmicas acerca dela durante o século XIX já foram exaustivamente comentadas no início desse trabalho, portanto abordaremos aqui os comentários de Barbosa (2005) a respeito disso.

Após a língua passar primeiro por um período fonético e posteriormente por um momento em que oscilava entre mais fonético e o etimológico, chegamos no período chamado pseudo-etimológico, que seria o século XIX, estado de língua que é o escopo de análise nesse trabalho.

Desse modo, uma forma de se estar ciente a novos padrões de época, era visto como importante suprimir as oscilações fonéticas nos textos, priorizando-se a volta de uso de expressões e formas gráficas das línguas de origem do português, o Latim e o Grego. Logo, o uso de uma escrita de forma etimológica seria visto como culto. Porém, com isso, surgiram diversas relatições não só etimologicamente corretas, como também errôneas, isto é, palavras com grafemas latinos ou gregos, mas que nem têm origem nessas línguas. Temos por exemplo *fallar* com origem *falare*, e *addiar* que vem da palavra *dia*, ambos vindos do latim, mas sem consoantes geminadas.

Há duas suposições que possibilitam a identificação de uma referencial objetivo para o controle desse grau de cultura (para assim, dessa forma, tentarmos diferenciar o que é culto, identificando o formal x informal). Com base nisso, o autor desenvolveu dois controles objetivos, nos quais se centravam basicamente em um maior uso de etimologizações, o que indicaria maior grau de cultura, e maior acerto indicaria maior escolaridade:

1. Quanto mais encontrarmos grafias latinas em um autógrafo oitocentista, mais seu autor teria estado em contato com os meios de transmissão de modelos de erudição, sejam escolares, de norma subjetiva ou de norma objetiva;
2. Se, e somente se, o uso da grafia for significativo – em quantidade e variedade consideráveis – quanto mais o referido redator mantiver um paralelo perfeito entre a forma latina oficial e o que escreve, maior seria o grau de conhecimento da norma culta de sua época via estudo – particular, público ou religiosa (BARBOSA, 2005, p. 31).

Como veremos melhor na próxima seção, isso acaba por se tornar não tão objetivo. Barbosa (2005) criou níveis, colocando o senador Ottoni no primeiro nível, como *erudito*; citar um nível intermediário, o de *informante culto*, como informante de grau intermediário; e

um terceiro como *informante de cultura mediana*, alguém alfabetizado, porém sem tanto contato com os meios de educação e cultura de sua época, ou seja, alguém como a avó Ottoni (BARBOSA, 2005, p. 38). Porém, não é comentado, quanto seria a taxa de acerto e de uso de etimologizações para cada nível.

Por fim, queremos verificar uma última coisa, que já acreditávamos ser possível e que agora acreditamos a veracidade com base no texto de Teyssier (1997). No texto sobre a história da nossa língua, o autor afirma que a dialetologia no Brasil é muito mais vertical do que horizontal. Ou seja, a diferença linguística se dá muito mais no âmbito sociocultural, do que no geográfico. Propomos, desse modo, utilizar essa metodologia em diferentes regiões do Brasil, para testar se esse é um método possível de ser utilizado com cartas oitocentistas de todo o país.

3.2 METODOLOGIA

Partindo-se do pressuposto de que textos-modelo do século XIX possam ser os jornais da época e de uma marca de registro ortográfico de prestígio era o uso de etimologizações, detalhamos a seguir como tornamos as hipóteses anteriores critérios objetivos para a classificação de grau de cultura dos redatores oitocentistas.

Uma vez que o determinante de grau de cultura é o uso e também o acerto das etimologizações em um texto algumas informações devem ter estratificadas: o número total de palavras do texto e o número total de palavras que foram etimologizadas. Os grafemas analisados foram consoantes geminadas (como *anno*, *telegramma*, *ella*), sequências mediais de obstruintes (*escripto*, *contractou*, *objecto*) e grafemas latinos (*telegrapharem*, *chronico*, *catastrophe*).

A partir do momento que temos o número total de palavras etimologizadas precisamos calcular quais delas são verdadeiras e quais delas foram etimologizações feitas erroneamente. Para isso, consultamos palavra por palavra no dicionário, para confirmar ou não sua origem. Assim, teremos três informações: número total de palavras no texto, número total de etimologizações e número total de etimologizações que foram feitas corretamente.

Dessa forma usamos a razão de palavras etimologizadas por total de palavras do texto para obtermos a taxa percentual do uso e a razão das palavras etimologizadas corretamente por total de palavras etimologizadas para obtermos a taxa de acerto do redator. Tendo essas

duas taxas conseguimos dividir os redatores em cinco graus de cultura: erudito, culto, culto-mediano, semi-cultos e mãos inábeis.

Como dito anteriormente, **erudito** seria alguém como o avô Ottoni, um senador, com alto grau de instrução e conhecimentos da norma culta escrita; um informante **culto** seria algo intermediário em relação a isso; um **culto-mediano** teria alfabetização, mas não grande contato com os meios de educação e cultura da época. Após isso, “fora e abaixo desse eixo de *cultural*”, como dito por Barbosa (2005), teríamos os **semicultos** e os **mãos inábeis**, termos que concordamos serem de difícil classificação.

Esses dois últimos termos são mais arriscados de serem utilizados, como o próprio autor reconhece. Ele afirmar que na própria bibliografia especializada **semicultos** é “confundido com a questão da inabilidade na escrita alfabética” (p. 38), porém, não explica o que englobaria esse rótulo. Já **mãos inábeis**, o último nível que teríamos, é de fato o mais complicado, pois está vinculado ao treinamento de escrita - ou seja, algo relacionado a crianças ou adultos recém alfabetizados. Esse termo, como já comentado anteriormente, é introduzido por Marquilhas (1996).

Tendo cinco níveis classificatórios, e analisando tanto nossos dados obtidos como os de Barbosa (2005) definimos quanto de uso e acerto um redator deve ter para entrar em cada um desses níveis: uso de etimologizações abaixo de 1% e taxa de acerto abaixo de 50% resulta em uma pessoa com **mãos inábeis**; uso de etimologizações em torno de 1% e taxa de acerto em torno de 50% seria **semi-culto**; uso de etimologizações em torno de 2% e taxa de acerto entre 60% a 80%, **culto mediano**; uso de etimologizações em torno de 3% e taxa de acerto em torno de 80%, **culto**; e por fim, uso de etimologizações acima de 4% e taxa de acerto acima de 90% resultaria em alguém **erudito**.

Como dito anteriormente, acreditamos nesse ser um método possível de ser aplicado em cartas de todo o Brasil, desde que do século XIX. Portanto contrastaremos cartas de redatores do sul do país com cartas do extremo oposto. Para tudo isso ser possível, é importante ressaltar que essa metodologia será feita em pessoas cujo perfil sociocultural se tem conhecimento, pois essa é a única maneira de comprovar que aquela pessoa realmente se enquadra em determinado nível sociocultural.

Com tudo isso em mente, conseguiríamos, em *corpuras* que não temos informações externas do redator, como nome, gênero, idade e região, algum dado sociocultural dele. Dessa forma, seria possível, por eliminação, e talvez mais algumas análises não tão objetivas,

classificar um texto como realmente formal e informal, e obter nosso tão desejado vernáculo. Tornando assim, parte dos *corpora* não utilizados possíveis de estudos e análises futuras.

3.3 AMOSTRA

Nossos testes foram realizados em 64 cartas pessoais, escritas por 25 diferentes pessoas. Foram utilizados dois diferentes *corpora*. O primeiro são cartas trocadas entre os familiares Prates de Castilhos, retiradas do Arquivo Histórico do Rio Grande do Sul (AHRs), do Museu Júlio de Castilhos e do Instituto Histórico e Geográfico do Rio Grande do Sul (IHGRGS). Já o segundo são cartas para Cícero Dantas Martins, Barão de Jeremoabo, retiradas do *Corpus* Eletrônico de Documentos Históricos do Sertão (CE-DOHS).

Formam o primeiro *corpus*: 54 cartas, trocadas entre 21 familiares, sendo 13 homens e 8 mulheres, de diferentes faixas etárias, nem todas confirmadas com exatidão. Contudo, a diversidade da idade dos informantes pode ser autenticada pelos diferentes níveis de parentesco em relação ao Júlio de Castilhos: irmãos, pais, sogra, cunhados, filhos, esposa, tios, avô e tio avô. A importância desse material se dá por existir vasta informação sobre os membros dessa família. Isso se dá pois Júlio de Castilhos, nascido no Rio Grande do Sul, em 1860, em Vila Rica, é uma personalidade muito conhecida na memória do Estado, e vinculou-se a aristocráticas famílias do sul do Estado.

Todas essas características já fazem o *corpus* ser interessante não só historicamente, mas também linguisticamente, por si só. Para nosso caso ele é interessante pelas informações complementares. Sabe-se que esse *corpus* é constituído por pessoas cultas, como o próprio Júlio, formado em Direito e que atuou como jornalista e político. Além do Júlio, sabemos que seus familiares tiveram estudos, mesmo que sem uma formação formal. Logicamente, a família por parte de sua esposa, também constituída por aristocratas, teve algum tipo de estudo, mesmo que menor. Sabemos dessas informações com base não só no conteúdo de algumas cartas, mas também de documentos encontrados nesses mesmos arquivos e museus.

As outras dez cartas, foram escritas por 3 homens e uma mulher, e fazem parte das cartas para Cícero Dantas Martins. Essas cartas foram escritas por sertanejos baianos, destinadas ao “coronel-barão” de Jeremoabo. Ele desempenhou diversos cargos políticos ao decorrer de sua vida, e conseguiu o título de barão quando fundou a primeira usina de açúcar do norte-nordeste, conhecida como “Fábrica Central do Bom Jardim”. Sendo essas cartas

destinadas a alguém de alta posição social, acreditamos se enquadrarem nos mesmos níveis que as cartas do primeiro *corpus*.

As cartas do primeiro corpus pertencem a uma amostra de um Projeto de Pesquisa, coordenado pela profa. Valéria Monaretto. Foram transcritas segundo normas semi-diplomáticas, segundo as quais se preserva a forma original de sua escrita, desmembrando-se palavras abreviadas, para compreensão do conteúdo.

O segundo corpus trata-se de uma edição fac-similar e semi-diplomática, disponibilizado-se diversas versões. A utilizada por nós foi a de texto simples (original), que assim como o corpus anterior preserva a escrita da época, com a diferença que neste as abreviações são mantidas como no original.

4 RESULTADOS

Apesar de toda a proposta girar em torno da utilização de um método objetivo para a análise de grau de cultura dos redatores, acreditamos ser importante algumas observações mais qualitativas. A relevância disso se dá justamente por ser um método ainda em desenvolvimento, portanto, precisamos comparar a sua veracidade da melhor forma possível. Sendo assim, comentaremos primeiro os dados do primeiro *corpus*, seguido por comparações com o trabalho de Barbosa (2005), seguido pelos dados do segundo e terminaremos por classificar de fato cada redator dentro de um dos graus de cultura propostos.

A tabela 1 mostra todos os resultados obtidos por nós na análise feita no primeiro *corpus*. A tabela foi feita elaborada por nós justamente com o propósito de compararmos resultados. Devido à grande quantidade de redatores, colocamos o parentesco em relação ao Júlio de Castilhos para facilitar a relação com ele e comparações feitas.

Tabela 1 – Relação de Formas Etimológicas em Cartas dos Familiares de Júlio de Castilhos (*corpus* 1)

Remetente	Uso de Etimologização	Taxa de Acerto	Taxa de Erro
Carlos Moreira (avô)	1% (2/201)	100% (2/2)	-
Carlos Prates (irmão)	1,38% (94/290)	75% (3/4)	25% (1/4)
Virgínia (cunhada)	1,47% (2/136)	100% (2/2)	-
F. Costa (cunhado)	1,54% (2/130)	100% (2/2)	-
Hermínio (cunhado)	1,59% (2/126)	100% (2/2)	-
Carolina (mãe)	1,68% (18/1071)	66,67% (12/18)	33,33% (6/18)
Honorina (esposa)	1,8% (10/555)	70% (7/10)	30% (3/10)
Etelvina (cunhada)	1,88% (3/160)	100% (3/3)	-
Joaquim (cunhado)	1,89% (7/371)	85,71% (6/7)	14,29% (1/7)
Fidêncio (tio)	1,92% (87/4526)	96,55 (84/87)	3,45% (3/87)
Feliciano (tio)	1,92% (2/104)	100% (2/2)	-
Ana Martins (sogra)	2,34 (37/1579)	97,3% (36/37)	2,7% (1/37)
Fidelis (tio)	2,64% (18/682)	55,56% (10/18)	44,44% (8/18)
Revoredo (cunhado)	2,9% (2/69)	50% (1/2)	50% (1/2)
Cassia Prates (irmã)	2,9% (59/2034)	89,83% (53/59)	10,71% (6/59)
Adelaide (irmã)	3,08% (22/714)	86,36% (19/22)	13,64% (3/22)
Francisco (pai)	3,45% (16/464)	56,25% (9/16)	43,75% (7/16)
Chiquinha (cunhada)	3,7% (9/243)	100% (9/9)	-
Dom Feliciano (tio avô)	3,83% (26/678)	84,62% (22/26)	15,38% (4/26)
Júlio	4,49% (174/3877)	92,53% (161/174)	7,47% (13/174)
Honório (filho)	5,45% (3/55)	66,67% (2/3)	33,33% (1/43)

(fonte: elaborada pelo autor)

A ordem de apresentação do uso de etimologizações nas tabelas 1 e 2 é crescente, para facilitar a análise. Com isso, o primeiro aspecto interessante de se observar, e talvez o primeiro que chame a atenção, é que, justamente os resultados que apresentam alguma inconsistência são daqueles redatores que tiveram baixo número de palavras. Seja por escreverem pouco em cada carta, ou por terem poucas cartas. Em geral é por escreverem pouco, pois Júlio, por exemplo, era uma pessoa que mesmo se analisássemos apenas uma carta já teríamos um número de palavras maior que duzentos.

Todos redatores que acertaram todas as etimologizações que escreveram, tiveram um número de palavras entre 100 e 200 palavras e o uso de etimologizações entre 2 e 3. É importante relatarmos que, em geral, essas palavras eram extremamente comuns, ou seja, uma palavra etimologizada tão usada que era de um conhecimento mais geral, ou todas as vezes a mesma palavra. Os pronomes *elle* e suas variações são os mais utilizados nesses casos, e mostram que apesar desses redatores utilizarem grafemas considerados de prestígio, talvez nem soubessem que ele fosse de fato um marcador da norma culta da época.

Os outros dois casos de baixo número de palavras e que fogem desse caso de 100% de acerto são o Revoredo, cunhado de Júlio, e o seu filho, Honório. Esses dois casos já se tornam um pouco mais difíceis de serem analisados objetivamente. Revoredo escreve, na verdade apenas uma palavra, e Honório duas. O que marcamos, respectivamente, como dois e três usos e um e dois acertos foi devido a uma palavra em específico, utilizada por ambos: a palavra *aceitar*.

Esse verbo vem do Latim *acceptare*, portanto, levando em consideração a norma culta da época, deveria ser grafado tanto com consoante geminada como também com encontro consonantal. Dessa forma, classificamos a palavra como um erro e um acerto. Isso pode de fato não ter sido a melhor solução, até porque se caso não tivéssemos feito-a, os dois redatores teriam se enquadrado perfeitamente aos demais casos de poucas palavras: taxa de acerto de 100% no uso de palavras etimologizadas.

O último caso de pouca quantidade de palavras e taxa de 100% de acerto, que aparece um pouco mais abaixo em nossa tabela 1, é o da Chiquinha, cunhada de Júlio. Chiquinha aparece com um número de uso de etimologizações bem expressivo, portanto, a porcentagem, em relação à quantidade de palavras que têm em suas cartas, faz com que ela apareça muito próxima de Júlio, segundo os dados expostos na tabela 1. O que nos faz estranhar essa informação é justamente ela ter uma taxa de uso similar a dele, porém um acerto que o supera.

Esse resultado reforça a problemática de se utilizar esse método com poucos dados, o que é natural em pesquisas em Língua Histórica.

O que ocorre é que Chiquinha usa 6 vezes a palavra *elle*, 2 a palavra *ella* e apenas uma vez a palavra *remetter*. De fato, estão todas grafadas corretamente para a época, porém temos aí uma variação de apenas dois casos de etimologizações. É bem frequente vermos os redatores, que já sabemos não serem tão cultos, terem alguma taxa de uso por causa do pronome *ele* e suas variações. O caso de Chiquinha se assemelha muito ao caso da avó Ottoni, que também teve uma taxa de acerto de 100%, porém um uso de apenas 1,29% como podemos ver na tabela abaixo.

Tabela 2 – Relação de Formas Etimológicas dos Avós Ottoni por Barbosa (2005)

Remetente	Uso de Etimologização	Taxa de Acerto	Taxa de Erro
Avó Ottoni	1,29%	100% (24/24)	-
Avô Ottoni	2,81% (204/7239)	92,15%	7,85%

(fonte: elaborada pelo autor)

É importante, para fins de conclusão, ressaltar que a avó Ottoni teve muito mais palavras analisadas, e muito mais ocorrências também. Mas, como Barbosa mesmo comenta “Apesar de acertar todos os latinismos que escreve (24 ocorrências), isso nada significa, pois são muito pouco variados e bem comuns os itens com grafia etimológica” (BARBOSA, 2005, p. 38).

Tirando esse caso atípico da Chiquinha, em geral temos toda a parte da família de Honorina, esposa de Júlio, na primeira metade da tabela, que também é onde mais agrupa as mulheres. Isso poderia indicar uma diferença cultural entre as duas famílias, sendo a de Júlio mais abastada que a de Honorina. E também poderia reforçar a questão de as mulheres não terem tanto acesso aos estudos na época.

Porém, um dado que nos surpreendeu, foi ver duas irmãs de Júlio, com uma taxa muito alta de uso e de acerto. Sabemos que não só Júlio como também seus irmãos tiveram um professor que lecionava em sua casa. Isso mostra, talvez, uma preocupação da família com os estudos sem tanta discriminação entre os gêneros, ou seja, uma mudança nos costumes da época.

A última observação que gostaria de fazer sobre os resultados do primeiro *corpus* é a dos casos de uma alta taxa de uso acompanhada de uma baixa taxa de acerto. São os casos de Fidelis e Francisco respectivamente tio e pai de Júlio. Trata-se de dois homens, de mesma

família, mesma geração, e mais velhos que Júlio. Com base nisso, ressalto a questão de prestígio dessa norma culta.

Era de conhecimento quase que geral, entre os redatores da época, a existência das etimologizações. Podemos supor o pronome *elle* como o mais comum, e levando em conta casos que praticamente só tem uma taxa de uso de etimologização por causa dele, nos perguntar se esse uso era de fato consciente. Saberiam esses redatores que existia um padrão de escrita, e esse padrão era marcado pelo uso de etimologizações? Ou essas pessoas grafavam dessa maneira apenas porque viam escrito assim?

Levantada essa questão, queremos contrapor usos de formas etimológicas com pessoas da época que não faziam isso de forma inconsciente. Pessoas pertencente à uma elite, que se preocupavam com algum tipo de prestígio social na forma de escrita, saberiam de que seria importante “andar de acordo” com um padrão de escrita valorizado na época.

Já dissemos anteriormente que essa tentativa de se usar formas latinas e gregas no registro do português do século XIX resultou nessas chamadas “falsas etimologizações”. Temos aqui o melhor exemplo para o fenômeno que acontecia na época: havia esse conhecimento de um prestígio de uma norma gráfica que era desejada mesmo por pessoas que não tinham tanto acesso à cultura e a estudos para usá-la corretamente, pois os casos de utilização são altos e de acertos, baixos.

A seguir, contrastaremos os resultados de nossos dois corpora, apesar de entendermos ser complicado uma comparação da análise de 21 diferentes pessoas com apenas 4. Como dito na seção anterior, esse segundo corpus trata de cartas de sertanejos baianos, destinadas a um Barão. Não foi feita uma pesquisa sobre esses redatores por nossa parte, então mesmo não sendo pessoas completamente desconhecidas, há pouca informação. Por conta do pouco material para análise, tivemos o cuidado de escolher a única mulher que faz parte das Cartas para Cícero Dantas Martins, para, pelo menos, tentarmos ter algum tipo de análise mais homogênea. Abaixo, a tabela elaborada a partir dos dados encontrados:

Tabela 3 – Relação de Formas Etimológicas em Cartas de sertanejos baianos a um Barão (*corpus 2*)

Remetente	Uso de Etimologização	Taxa de Acerto	Taxa de Erro
Mariana	3,23% (10/310)	100% (10/10)	-
Alexandre	4,81% (73/1518)	78,05% (57/73)	21,92% (16/73)
Annibal	5,26% (16/304)	100% (16/16)	-
Amaro	5,28% (18/341)	77,78% (14/18)	22,22 (4/18)

(fonte: elaborada pelo autor)

São muitas as possíveis primeiras impressões dessa tabela, apesar de extremamente menor que a primeira. Assim como a tabela 1 temos aqui casos de poucas palavras com totalidade de acertos. Mas ao mesmo tempo temos o caso de Amaro, que mesmo tendo pouco material, teve as taxas de acerto e erro mais equilibradas. De qualquer modo, mesmo esses casos de 100% de acerto, encontramos características diferentes dos casos anteriores.

Como explicamos, havia esses redatores com um uso relativamente baixo de etimologizações se comparado aos demais, porém com uma taxa de acerto perfeita. Os redatores que, apesar de se destacarem pela elevada taxa de uso, pecavam no quesito acerto. Temos aqui duas novas situações.

Primeiro que esses casos de total acerto têm uma taxa excelente de uso. Mesmo Mariana, que é a que menos usa entre os remetentes analisados, tem uma taxa de uso maior que grande maioria dos redatores do *corpus* da família do Júlio (ficando atrás de apenas cinco pessoas). Annibal parece mais atípico ainda, pois apresenta uma taxa de uso maior que o próprio Júlio de Castilhos, ficando pra trás apenas de Honório (que já vimos ser um caso complicado a se analisar).

Porém, o que mais surpreende nesses dois casos, é que, mesmo acertando todas as palavras utilizadas, não foram palavras repetidas como o caso de *Chiquinha*, por exemplo (do *corpus* 1 de Júlio de Castilhos). *Marina* repete o uso de apenas duas palavras em duas cartas, e *Annibal* repete três em uma carta - uma delas o pronome *elle*, que foi utilizado cinco vezes, o que fez com que se subisse a taxa de uso de etimologizações.

Fora isso, aos redatores desse *corpus* 2, além de usarem um número diversificado de palavras, mesmo em cartas curtas, ambos usam palavras não tão comuns. *Annibal* utilizou a palavra *paranypho* corretamente, que apesar de não ser uma palavra de origem latina, mas sim grega, está etimologicamente correta. *Marina* usa sim palavras bem comuns como o pronome *elle* e *annos*, muito frequentes no primeiro *corpus*, mas também utiliza *escripto* e *approvadas* de maneira correta, de acordo os padrões da época.

Dessa maneira, nenhum desses dois remetentes podem ser enquadrados junto com os casos anteriores. Nem ao caso de pouco uso e muito acerto, nem ao caso de muito uso e pouco acerto. Mesmo assim, podemos supor alguma informação sociocultural. *Amaro* erra apenas duas palavras: *fallar* e *demittir*. Nenhuma delas utilizada por *Mariana*.

Dentre essas duas palavras, poderíamos considerar a palavra *fallar* uma palavra comum, e com isso gostaria de levar nosso leitor a uma ideia que surgiu ao observarmos outro

redator, desse modo comentamos de novo sobre a consciência do redator. A palavra *falar* não é escrita erroneamente apenas por ele, mas por várias pessoas, inclusive por Júlio de Castilhos. Portanto, o ponto que gostaríamos de trazer é: o quanto esse erro era, de fato, um erro?

Essa questão será melhor desenvolvida em nossas conclusões, portanto, aqui, o que queremos comentar, é: provavelmente *Mariana* fosse mais familiarizada com a norma culta, porém não tivesse tanto acesso aos meios de erudição da época. Já *Annibal* cometeria erros por estar mais exposto a esse meio, porém não ter tanto estudo formal sobre a norma. Ele também poderia ter mais conhecimento ou maior preocupação em relação a ter que usar etimologizações para um texto se enquadrar aos padrões da época.

Por fim, temos os casos de Alexandre e Amaro. Não há tanta discrepância entre ambos. Os dois tem praticamente a mesma taxa de erro, acerto e de uso muito parecido. O comentário a ser feito nesse caso enquadra também o remetente, já comentado, Annibal. Surpreendemo-nos com uma taxa tão alta de uso de etimologizações, que não é resultado de mera repetição de vocábulos comuns. Todos nesse *corpus* utilizam uma variedade muito expressiva de palavras, algo que não é visto na maioria das cartas dos familiares de Júlio de Castilhos.

Não fizemos nenhum comentário desse último *corpus* comparando diretamente com os resultados dos avós Ottoni pelo simples motivo de que os avós já se distanciavam do primeiro *corpus* por uso de etimologias. Nesse, então, se distanciam ainda mais. Porém, deve ser ressaltado que essas observações foram feitas graças as constatações feitas entre as duas primeiras Tabelas. Deixamos, então, ao fim da seção, como uma síntese conclusiva, o quadro 1 com uma classificação, conforme os cinco níveis propostos, com nossos dois *corpora* e o resultado de Barbosa (2005).

Quadro 1 – Classificação de Graus de Erudição dos *Corpora* Analisados

Grau de Erudição	Corpus 1	Corpus 2	Avós Ottoni*
mãos inábeis Uso de etimologizações abaixo de 1% e taxa de acerto abaixo de 50%			
semi-cultos Uso de etimologizações em torno de 1% e taxa de acerto em torno de 50%			
culto mediano Uso de etimologizações em torno de 2% e taxa de acerto entre 60% a 80%	Ana Martins; Joaquim; Honoria; Carolina.		Avó Ottoni

Grau de Erudição	Corpus 1	Corpus 2	Avós Ottoni*
culto Uso de etimologizações em torno de 3% e taxa de acerto em torno de 80%	Dom Feliciano; Adelaide; Cassia	Mariana	
erudito Uso de etimologizações acima de 4% e taxa de acerto acima de 90%	Júlio de Castilhos	Alexandre; Anibal; Amaro	Avô Ottoni

*classificação de acordo com Barbosa (2005)
(fonte: elaborado pelo autor)

Antes de finalmente concluirmos esse trabalho, eis algumas explicações de escolhas tomadas no preenchimento desse Quadro 1, que podem trazer algum estranhamento. Primeiro, no *corpus* da família de Júlio, não colocamos os casos que fogem demasiadamente dos padrões de nossa classificação, ou seja, os casos de 100% de acerto e os casos de muita irregularidade, por terem muito uso acompanhado de pouco acerto.

Porém, no caso do segundo *corpus*, como analisamos menos remetentes, e levando em conta esses apresentarem uma grande diversidade lexical, fato que dá mais crédito a uma taxa alta de uso de etimologizações, colocamos todos eles em uma classificação, principalmente de acordo com o uso de etimologizações. Por fim, colocamos os avós Ottoni, em nosso quadro 1, de acordo com a classificação de Barbosa (2005). Segundo nossa classificação, o avô Ottoni se enquadraria como culto, e sua esposa, a avó Ottoni, entraria entre nossos casos de muito acerto, porém pouco uso e, principalmente, pouca diversidade lexical.

5 CONCLUSÃO

Neste trabalho tentamos de diversas formas tanto situar o leitor sobre alguns dos problemas encontrados pelos pesquisadores de Linguística Histórica, como também apresentar nossa linha de pensamento. Ou seja, tentamos exemplificar porque existem esses problemas e porque são tão importantes de serem superados. Apesar de sabermos que essa tarefa é extremamente complexa, deixamos aqui algumas conclusões, com o propósito não só de dar um fechamento a esse trabalho, mas principalmente, dar novas ideias para possíveis próximas pesquisas que buscam desenvolver métodos investigativos que ajudem na área de Linguística Histórica.

Para isso, começamos justamente explicando sobre a norma culta do século XIX, ou melhor, na dificuldade de se classificar algo como norma culta para essa época. Atualmente, ou até mesmo em outros períodos, poderíamos simplesmente começar essa questão pela gramática, vendo o que era prescrito naquela época. Porém, países como o Brasil, colonizados a uma longa distância de seu país colonizador e repleto de outras culturas, portanto, de outras línguas, tiveram problemas ao formar uma só norma que fosse aceita e utilizada por todos.

Vimos a dificuldade que o Brasil teve de constituir a sua própria língua. A própria Língua Portuguesa passou por muitas mudanças, e teve contato com diversos outros idiomas. É interessante de se ressaltar que justo na época aqui estudada, tivemos tantas polêmicas em torno desse assunto. Todos os intelectuais pareciam ter uma opinião própria, cada um seguindo uma linha e considerando ela a correta. Debates sobre seguir os padrões do português europeu não foram traçados. Havia constante dúvida com grafia, posições pronominais e até mesmo entre uso analítico ou sintético do gerúndio, dentre outros aspectos linguísticos.

Tudo isso foi feito ignorando a língua da maioria, a língua cotidiana, falada pela população, ou seja, ignorando as mudanças que o idioma sofreu ao chegar aqui no Brasil. Essas pessoas não tinham voz. Não foi sob elas que pairou a discussão que perdura até hoje sobre não sabermos falar nossa língua. Essa crítica era feita sob as pessoas que, mesmo sem serem grandes intelectuais as que discutiam o futuro da Língua, tinham algum *status* diferencial: elas sabiam ler, escreviam, e por vezes, não sabiam que deviam ou não tinham conhecimento dessa norma culta escrita.

Esse contexto por si só já atrapalha o pesquisador que pretende se arriscar por estudar material escrito dessa época. Por isso, buscamos autores que encontraram algum meio

de driblar essa falta de norma e buscarem, eles mesmos, algum ditador de norma escrita para esse período. Como vimos, os pesquisadores utilizam os jornais para ter alguma noção dos padrões da época, o que se mostrou útil e eficaz.

Com esse panorama conseguimos dar concretude para a importância de nossos objetivos, e conseqüentemente de nossas hipóteses, uma vez que a tivemos com base não só de textos sobre a formação de nossa língua como também em trabalhos de outros pesquisadores. Ao decorrer do nosso trabalho, no entanto, nossa tarefa de encontrar um determinante objetivo para classificação de grau de cultura para redatores do século XIX, foi se mostrando mais complexa que o imaginado.

Analisando nossos primeiros dados, logo concluímos, como já havia sido imaginado antes, que por se tratar de um teste objetivo, ele seria inconsistente nos casos de pouco material, e com pouco material, queremos dizer redatores com um baixo número de palavras a serem consideradas. De qualquer modo, seguindo uma classificação feita com base no redator com resultados mais altos, ficamos satisfeitos por não classificar ninguém nos níveis mais baixos de grau de cultura.

Esse resultado corrobora de fato com o trabalho de Barbosa (2005). Levando em consideração a similaridade de perfis do avô Ottoni e de Júlio de Castilhos, o teste se mostraria eficaz caso os classificasse, através desse método, como dois redatores de mesmo grau de cultura. Ambos se assemelham pelo uso e pelo acerto das etimologizações escritas, e isso torna mais curioso o caso da avó Ottoni, e dos redatores que não classificamos no quadro 1.

No nosso caso, julgamos talvez erroneamente, esses redatores com uso perfeito das etimologizações. Levando em consideração que os mais cultos, ou seja, os que mais utilizam etimologizações, erram certas grafias, supõem-se que essa taxa de acerto também vá diminuindo à medida que o uso diminui. No momento que vimos isso não ocorrer justamente nos casos de pessoas com pouco material analisado, consideramos essa a causa de tal discrepância.

Isso entra, primeiramente, em conflito direto com o caso da avó Ottoni. O caso dela não se enquadra em poucas palavras analisadas. Porém, esse fato foi deixado de lado pelo fato dela apresentar ou similaridade com o nosso caso: baixa diversidade de palavras. Apenas percebemos tratar de talvez outro problema, quando nos deparamos com os casos do segundo

corpus: redatores com a mesma característica de todas etimologizações utilizadas corretamente, mas agora, com duas diferentes características.

A primeira em relação a ambos os *corpora*: pessoas com alta taxa de uso de etimologizações e com diversidade de palavras. A segunda: redatores com um número baixo de palavras analisadas. Apesar de isso trazer mais indagações, podemos sim concluir algumas coisas, começando pelo fato de que temos que continuar considerando esse método como uma investigação em fase inicial, portanto, mantê-la sob suspeita, como o próprio Barbosa diz nas considerações finais de sua pesquisa de 2005.

Mas acreditamos ter achado aí, a chave para o aperfeiçoamento dessa metodologia. É necessário, sim, considerarmos a diversidade de palavras encontradas num texto. Levar isso em consideração parece ser, de acordo com os resultados, até mais importante que a necessidade de utilizar apenas redatores com grande quantidade de palavras escritas para aplicação dessa metodologia. E com isso, mantemos e reforçamos sim nosso objetivo.

Dessa forma, admitimos que essa metodologia, apesar de mais desenvolvida que a inicialmente proposta por Barbosa (2005) ainda não está aperfeiçoada. Esse trabalho não nos permite ir para essa próxima etapa de aperfeiçoamento da metodologia. Porém, frisamos aqui, estamos cada vez mais perto de conseguir de fato evitarmos análises extremamente longas e quantitativas e conseguirmos uma aplicação em massa.

Confirmamos essa possibilidade através de conceito que não foi abordado nesse trabalho, mas que supre essa necessidade. De acordo com a definição simples, encontrada no site do Projeto PorPopular *riqueza lexical* “[...] uma medida estatística [...] que se estabelece entre o número de palavras repetidas e diferentes de um texto e o número total de palavras nele encontradas” (FINATTO, 2010). Ou seja, através da razão de *types* (número de palavras diferentes no texto) por *tokens* (número total de palavras do texto) obtemos a diversidade lexical de um texto, sendo que quanto maior o *types* maior a variedade do vocabulário, quanto menor, mais repetitivo o texto é.

Por fim, após essas diversas conclusões, frisamos que, apesar de termos conseguido classificar pessoas de três regiões em uma mesma tabela, através de um método objetivo, apesar de ainda imperfeito, mais testes devem ser feitos. O segundo *corpus* não tem tantos redatores, nem número expressivo de palavras utilizadas pela maioria, nem abrange tanta diversidade de perfis sociais como o primeiro. A primeiro olhar sim, esse teste poderia ser

aplicado em diferentes regiões do Brasil. Mas o fato de esses redatores terem uma taxa tão alta de uso de etimologizações não pode ser ignorado.

Esperamos que nossa pesquisa gere mais frutos, uma vez que, mesmo não tendo cumprido perfeitamente com todas suas propostas, ele teve ótimos resultados. Não só conseguimos evoluir uma ideia inicial, como também ter uma ideia de quais são os próximos passos. Ter a oportunidade de compartilhar esses dados é de suma importância, uma vez que possibilita que a pesquisa siga em frente, ou simplesmente dê novas ideias a quem for lê-la.

REFERÊNCIAS BIBLIOGRÁFICAS

- BARBOSA, A. G. **Tratamento dos Corpora de sincronias passadas da Língua Portuguesa no Brasil**: recortes grafológicos e linguísticos. Rio de Janeiro: IN-FÓLIO, 2005.
- BUENO, F. S. **A Formação Histórica da Língua Portuguesa**. 2. ed. Rio de Janeiro: Livraria Acadêmica, 1958.
- CÂMARA JÚNIOR, J. M. **História e Estrutura da Língua Portuguesa**. Rio de Janeiro: Padrão - livraria editora LTDA., 1975.
- FARACO, C. A. **Linguística Histórica**: uma introdução ao estudo da história das línguas. 2. ed. São Paulo: Parábola Editorial, 2005.
- FINATTO, M. J. B. **Riqueza Lexical**. Porto Alegre, 2010. Disponível em: <http://www.ufrgs.br/textecc/porlexbras/porpopular/riqueza_lexical.php>. Acesso em: 10 dez. 2018.
- LOPES, C. R. S. **A História da Família Ottoni nas linhas e entrelinhas**. Rio de Janeiro: IN-FÓLIO, 2005.
- LOPES, C. R. S.; MARCOTULIO, L. L.; RUMEU, M. C. B.; LIMA, A. X. Reflexões Metodológicas para a Análise Sociocultural de Redatores em Corpora Históricos. **Gragoatá**, Niterói, v. 15, n. 29, p. 239-253, 2010.
- MARQUILHAS, R. A faculdade das letras: leitura e escrita em Portugal no século XVII. **Universidade de Lisboa**, Lisboa, v.2, p. 239-257, 1996.
- PAGOTTO, Norma e Condescendência; Ciência e Pureza. **Línguas e Instrumentos Linguísticos**, Campinas, n. 2, p. 49-68, 1999.
- _____. Rui Barbosa e a Crise Normativa Brasileira, 2011.
- SÍLVIA, E. N.; LOPES, C. R. S. O perfil sociolinguístico de um casal não ilustre: uma análise grafemática através da edição de cartas particulares. **Confluência**, Rio de Janeiro, v. 1, p. 78-104, 2013.
- TEYSSIER, P. **História da Língua Portuguesa**. Tradução de Celso Cunha. 1. ed. São Paulo: Martins Fontes, 1997.