

PESQUISAS SOBRE O PORTUGUÊS PAUTADAS EM CORPORA

RESEARCH ON THE PORTUGUESE IN CORPORA

Daniel Soares da Costa

UNESP - Araraquara

Juliana Bertucci Barbosa

UFTM

Talita de Cássia Marine

UFU

Niguelme Cardoso Arruda

FAFICH

RESUMO: Neste artigo, apresentamos a investigação de quatro fenômenos linguísticos da Língua Portuguesa em diferentes sincronias. Essas pesquisas abordam a montagem, a descrição e a análise de *corpora* para a constituição de bancos de dados que dão suporte à descrição linguística.

PALAVRAS-CHAVES: *corpus*; variação e mudança linguísticas; português.

ABSTRACT: This article presents an investigation of four linguistic phenomena of the Portuguese in different synchronicities. These research are about construction, description and analysis of *corpora* for the creation of databases that support linguistic description.

KEYWORDS: *corpus*; variation and language change, Portuguese.

1. Considerações Iniciais

O objetivo deste artigo é congrega pesquisas de diferentes áreas dos estudos linguísticos, especialmente com foco no estudo da Língua Portuguesa, que trabalham com a montagem, descrição e análise de *corpora* para a constituição de bancos de dados que dão suporte à descrição linguística, seja em estudos sincrônicos ou diacrônicos. Independentemente do gênero textual, ou da modalidade da língua (escrita ou falada), a coleta de dados fornece um suporte empírico para os pesquisadores, permitindo a aplicação de reflexões teóricas sobre fatos do uso real da língua.

Trata-se de um tema bastante abrangente, que abre a possibilidade de colocar lado a lado tanto estudos de caráter formalista quanto funcionalista, já que a montagem de *corpus* permeia grande parte dos estudos nessas duas áreas. A montagem de *corpus* requer o uso de tecnologia, sendo indispensável, na atualidade, a utilização de, no

mínimo, um computador para registro de dados. Se o foco da pesquisa for a língua falada, também são necessários recursos para a gravação de áudio, tais como estúdios com isolamento acústico, além de softwares para a análise acústica dos sons, principalmente em estudos de fonética. Porém, outros recursos podem ser utilizados para se estudar a língua falada. Um deles é a utilização de programas de TV ou rádio ao vivo que, de certa forma, fornecem dados espontâneos do uso da língua, principalmente em programas de entrevista com auditório. Por outro lado, quando os estudos são voltados para períodos passados da língua, principalmente em relação a períodos em que ainda não havia a tecnologia de gravação de áudio, os pesquisadores devem lançar mão do recurso a textos escritos, por meio dos quais é possível buscar pistas do funcionamento da língua na época escolhida. Nesse caso, torna-se imprescindível o acesso a documentos escritos antigos, os quais nem sempre foram preservados de maneira adequada e acabaram, muitas vezes, deteriorados, o que dificulta a leitura por parte dos pesquisadores.

Sendo assim, em alguns casos, além do conhecimento da sua própria área, a Linguística, o pesquisador necessita de conhecimentos de Paleografia, para o manuseio e a preservação de documentos escritos antigos e, em alguns casos, um treinamento de leitura, quando lida com documentos com estilos diferentes de escrita como, por exemplo, a escrita gótica em documentos medievais. Por outro lado, além do conhecimento sobre o uso da tecnologia a favor da pesquisa, de conhecimentos em outras áreas, no caso de levantamento de dados que envolvam documentos antigos, faz-se necessário também conhecer o tipo e o gênero textual que constituirá o corpus do estudo, pois cada gênero apresenta peculiaridades nos seus aspectos estruturais, nos temas tratados, além da questão do grau de formalidade, o que interfere na análise e interpretação dos dados coletados.

2. Estudos linguísticos pautados em *corpus*

Conforme as autoras Berlinck, Barbosa e Marine (2008) afirmam, estudos baseados em *corpus* – tanto oral quanto escrito - proporcionam novas descrições e hipóteses teóricas, fortalecendo e tornando mais consistentes as pesquisas linguísticas desenvolvidas. Isso porque sua utilização proporciona a realização de descrições linguísticas de base empírica, além de permitir a reflexão sobre questões teóricas fundamentadas em contextos de usos reais da língua.

Ainda vale ressaltar que um dos mais intrincados desafios com o qual o pesquisador da língua se depara é o das fontes para o seu estudo em diferentes épocas. Por ter como objetivo o registro de um determinado estado passado de língua (sincronia no passado) ou o traçado das mudanças linguísticas ao longo de extensos períodos de tempo (diacronia), o pesquisador trabalha, inevitavelmente, com registros escritos em épocas passadas (BERLINCK, BARBOSA E MARINE, 2008).

Baseadas no modelo teórico-metodológico laboviano (LABOV, 1972, 1994; 2001), as pesquisadoras afirmam que, ao se optar pelo trabalho com textos escritos não contemporâneos, a tendência é a busca pelo texto que represente um menor grau de formalidade e/ou que mais se aproxime da fala. Entretanto, alertam que tal escolha pode acarretar a exclusão de textos tradicionalmente considerados mais formais, tais como editoriais, documentos oficiais, textos pertencentes à literatura oratória, entre outros, mesmo diante do fato de que um linguista não deve pré-avaliar um texto como formal e informal a partir de uma perspectiva atual que ele possua acerca de determinado gênero.

As autoras propõem ainda que, para auxiliar no estudo da variação e mudança linguística, é preciso ter um maior critério na seleção e avaliação das fontes, de modo que possibilite uma ampliação das possibilidades de manifestações linguísticas investigadas nos estudos de variação e mudança.

Assim, ao se optar pela pesquisa linguística baseada em *corpus*, cabe ao linguista conhecer o gênero textual com que irá trabalhar durante sua pesquisa, visto que cada gênero textual apresenta peculiaridades que implicam não apenas em aspectos estruturais caracterizadores, mas também em tipos de temas mais recorrentes, grau de formalidade, estilo, intenção, dentre outros. Ao conhecer tais características, o pesquisador acaba se vendo diante de um caminho mais adequado para o tratamento do texto utilizado como *corpus*, seja oral ou escrito, o que, por sua vez, proporcionar-lhe-á uma análise mais exata do seu objeto de estudo.

3. Processos morfofonológicos e flexão nominal no português arcaico por meio das Cantigas de Santa Maria

Nesta seção, apresentamos resultados parciais do projeto de pesquisa intitulado “Processos morfofonológicos na flexão nominal: estudo comparativo entre o português arcaico e o português brasileiro atual”, que tem como objetivo fazer um mapeamento dos processos morfofonológicos desencadeados pela flexão nominal do português de modo a

se estabelecer uma análise comparativa entre duas sincronias dessa língua: o Português Arcaico (de agora em diante PA) - séculos XII e XIII - e o Português Brasileiro (de agora em diante PB) - séculos XX e XXI.

A Morfofonologia é o estudo das variações fonológicas que sofrem os elementos morfológicos - radicais, afixos, vogal temática e desinências. Um processo morfofonológico ocorre quando uma regra fonológica é desencadeada por um processo morfológico. Podemos citar, como exemplos de processos morfofonológicos, a **alternância vocálica**, que ocorre na formação do plural de palavras como “formoso > formosos”, em que há uma abertura da vogal “o” da sílaba “mo” como reforço para a indicação de plural; a **supressão** da vogal temática nominal, como em “lobo > lobinho”; a **crase**, que é a fusão de duas vogais idênticas, como “teme + ia > /temi + ia/ > /temia/”; e a **assimilação** de traços, como em “canta + i > cantei”, em que a vogal baixa /a/ (vogal temática verbal) assimila o traço de altura da vogal alta /i/ (vogal inicial do morfema modo temporal de Imperfeito do Indicativo), tornando-se uma vogal média-alta /e/. Processos morfofonológicos podem ser desencadeados tanto pela flexão quanto por processos de formação de palavras, tais como derivação e composição.

Por meio desse estudo, pretendemos estabelecer um paralelo entre as variações fonológicas sofridas pelos morfemas no PA e no PB atual, no que diz respeito à flexão nominal, a fim de observar quais são os processos morfofonológicos mais recorrentes em cada uma dessas fases da língua e também verificar se os processos que ocorrem no PB atual já ocorriam no PA.

Para o desenvolvimento desse projeto, utilizaremos dois *corpora* representativos de cada período da língua portuguesa que pretendemos comparar. O *corpus* que sustentará a análise dos processos morfofonológicos desencadeados pela flexão nominal no PA será constituído a partir do texto das *Cantigas de Santa Maria* (de agora em diante CSM) de Afonso X, o Rei Sábio de Leão e Castela.

O segundo tipo de *corpus*, que sustentará a análise do PB dos séculos XX e XXI, será composto por um recorte do banco de dados do Laboratório de Estudos Lexicográficos da UNESP (LabLEX) para o PB.

O primeiro passo para a organização dos *corpora* para o procedimento da análise dos processos morfofonológicos desencadeados pela flexão nominal é a separação, por classes morfológicas, das palavras que sofrem a flexão de número e gênero, no banco de dados fornecido pelas CSM para o PA. Dessa forma, podemos dizer que aparecerão,

nessas condições, **substantivos**, **adjetivos**, **artigos**, alguns **numerais** e alguns **pronomes**.

A partir dessa separação, encaminhar-nos-emos para a identificação dos tipos de processos morfofonológicos que apareceram nesse contexto linguístico, procedendo também com a quantificação de palavras em relação a cada processo morfofonológico identificado.

Até o momento, foram coletados dados das 10 primeiras Cantigas de Santa Maria, que estão relacionados na tabela 1 a seguir.

Tabela 1. Quantificação Geral

Substantivo	782	37,57%
Adjetivo	170	8,16%
Particípio	63	3,02%
Artigo	294	14,12%
Preposição	1?	0,04%
Preposição mais artigo	124	5,95%
Advérbio mais artigo	1	0,04%
Pronome mais artigo	2	0,09%
Pronome	618	29,69%
Preposição mais pronome	19	0,91%
Advérbio mais pronome	1	0,04%
Advérbio	1?	0,04%
Numeral	5	0,24%
Total	2081	100%

Na tabela acima, estão contempladas também a classe da preposição e do advérbio, que são palavras invariáveis no PB atual. No entanto, as palavras encontradas no PA são estranhas à língua atual e possuem elementos fonéticos que nos fazem levantar a possibilidade de que sejam oriundos de processos flexionais. Porém, somente uma análise mais acurada do comportamento de tais palavras no PA e também a análise da sua etimologia é que poderão fornecer informações para afirmações mais contundentes a esse respeito. Por isso o ponto de interrogação em frente ao número de vezes em que a palavra apareceu.

Dentre as palavras coletadas, nem todas se apresentaram com flexão nominal. Em relação à flexão de número, podemos afirmar que a regra geral, assim como no PB atual, se dá pelo acréscimo do morfe -s a palavras terminadas em vogal, uma vez que, no corpus, mais de 86% das palavras flexionadas em número se encaixam nessa regra. As palavras terminadas em consoante, que apresentaram flexão de número, representam 4,87% e recebem o acréscimo do alomorfe -es.

Quanto às alomorfas, encontramos um possível caso de alomorfe Ø, parecido com casos como “o ourives > os ourives; o lápis > os lápis” do PB atual. Trata-se da palavra “lais”, que aparece nos contextos “...un seu lais senpre dizia...”, “...O lais que ele cantava...”. Pela configuração fônica da palavra, podemos levantar a hipótese de que seria um caso igual aos citados, porém só o levantamento de mais dados poderá confirmar essa hipótese.

Também encontramos um possível caso de alternância vocálica submorfêmica, como acontece em palavras do PB como “formoso > formosos; olho > olhos”, em que há a abertura da vogal “o” /formɔzus/ e /ɔlus/. A nossa suspeita é a de que essa alternância vocálica não acontecia no PA. Essa hipótese parte da análise de duas palavras que aparecem rimando entre si nos seguintes versos da Cantiga 009:

Log' enton a dona, | chorando dos **ollos**,
muito lle rogava | que per y tornasse,
des que el ouvesse | fito-los **gẽollos**
ant' o San Sepulcro | e en el beijasse.

Como a vogal “o” da palavra “gẽollos” aparece antecédida por uma vogal nasalizada, provavelmente essa nasalização impede a abertura da vogal, sendo, portanto, um caso de “o” fechado. Uma vez que essa palavra está rimando com a palavra “ollos”, no primeiro verso, provavelmente a abertura do primeiro “o” de “ollos” também não acontecia no plural, sendo, portanto, fechado.

Em relação a palavras terminadas em -ão, a única forma de plural encontrada foi com o acréscimo de -s apenas, como em “crischão > crischãos, irmão > irmãos, mão > mãos, vilão > vilãos”. Os plurais terminados em -ões têm, como forma de singular, palavras terminadas em -on, como em “don > does, leon > leões, varon > varões”.

Em relação à flexão de gênero, no PA, podemos dizer que a regra geral também se dá pelo acréscimo do morfe -a à forma do masculino, assim como no PB. Alguns casos que se diferem um pouco estão relacionados às alomorfas. Como nos casos de palavras terminadas em nasal que, em PB, ou perdem a nasalidade ou ela passa para a sílaba seguinte na posição de *onset*, em um processo de ressilabificação, como em “leão > leoa (perda da nasalização), valentão > valentona (ressilabificação)”. No PA, a nasalidade parece ser mantida, o que pode ser comprovado pelos exemplos “bon > bõa, un > hũa”.

Outro caso interessante é o caso de ditongação com alternância vocálica, que acontece, em PB, em palavras como “europeu > europeia”. No entanto, no PA, esse fenômeno ainda não acontecia, como podemos perceber pelo exemplo “judeu > judea”.

Para finalizar, gostaríamos de dizer que pretendemos chegar até a análise das cem primeiras CSM. Depois, faremos o levantamento dos dados do PB atual, de modo a termos uma quantidade de dados parecida com a quantidade levantada para o PA. Após o levantamento e organização dos dados, daremos início às análises, primeiramente tendo como bases teóricas a Fonologia Lexical e a Teoria de Geometria de Traços, e, posteriormente, a Teoria da Otimalidade, de modo a podermos compará-la às teorias não-lineares, a fim de observarmos a adequação dessas teorias em relação à análise dos processos morfofonológicos desencadeados pela flexão de número nessas duas línguas.

4. Variação linguística no português mineiro: o emprego das formas verbais do indicativo nas páginas de jornais antigos de Uberaba-MG

Esta seção apresenta outra investigação linguística baseada em *corpus*, do projeto intitulado “Variação linguística nas páginas de jornais uberabenses: um estudo de formas verbais” que visa, principalmente: (a) estudar das formas verbais do modo Indicativo, em textos jornalísticos escritos na cidade de Uberaba, séculos XIX e XX; (b) investigar questões de natureza conceptual e empírica – análise estatística dos dados –, seguindo os objetivos do modelo variacionista laboviano, abordando os processos de variação e mudanças sofridos pelas formas verbais do modo Indicativo; (c) a associar dos estudos de variação e mudança nos usos das formas verbais ao contexto sócio-histórico do período; (d) a contribuir na organização de bancos de dados com textos escritos de sincronias passadas do Português Mineiro (PM), mais especificamente, do português de Uberaba e região.

Este projeto, ainda está em andamento, o que tentamos evidenciar aqui é a pertinência de se estudar as formas verbais e seus valores semânticos em *corpus*. O *corpus* que serve de fonte para a análise das formas verbais será composto por jornais publicados na cidade de Uberaba (séculos XIX e XX). Entendemos como uma importante fonte para a organização de um *corpus* a chamada língua da imprensa. Esse tipo de arquivo como fonte de dados para estudos de variação é pertinente pelo fato de estudarmos uma sincronia passada que não dispõe de fontes de língua falada e por

permitir um espaço de discussão entre norma gramatical, variação e mudança. Os jornais visam compor um *continuum* de formalidade (editorias, anúncios, cartas etc).

4.1. Por que estudar formais verbais do modo Indicativo?

Partindo da concepção de língua heterogênea, consideramos que o Português Mineiro (PM), como todo sistema linguístico, comporta em seu interior a variação e, assim, está continuamente em processo de mudança. Vários fenômenos de variação no português brasileiro têm merecido a atenção dos estudiosos: o preenchimento da posição do sujeito por pronomes plenos, a manutenção dos sintagmas interrogativos *in situ*, o aumento na ocorrência dos objetos nulos, na sintaxe; a perda de sufixos e flexões, com o conseqüente surgimento de formas verbais analíticas a partir da gramaticalização de verbos auxiliares, e com o recurso mais frequente à recategorização de itens, na morfologia (cf. SILVA, 1997; DUARTE, 1996; CYRINO, 1997; LONGO, 2000, BARBOSA, 2003, 2008, SILVA, 2008 entre outros).

No sistema verbal do português também podemos encontrar variações e/ou mudanças. No que diz respeito à flexão verbal, os afixos característicos de segunda pessoa têm rendimento praticamente nulo, e foram substituídos em quase todas as variedades de linguagem por morfemas originalmente empregados para a terceira.

Também quanto aos modos e tempos houve rearranjos: as mais estudadas são as flexões de futuro do presente, que atualmente só ocorrem em contextos muito específicos, estando restritas, de acordo com pesquisas recentes (cf. SILVA, 1997; MOTA, 1998; BARBOSA, 1999; BARBOSA E LONGO, 2001) a textos altamente formais, de teor preditivo ou injuntivo. Nos demais casos, o futuro flexionado é substituído pelo presente do indicativo ou por perífrases, das quais a mais gramaticalizada é *ir + infinitivo*. Ao mesmo tempo, observa-se a penetração de uma forma perifrástica composta por *estar + gerúndio*, que ainda não mereceu estudos mais aprofundados, mas que provavelmente tem força modal, além de valor prospectivo, como se pode concluir por ocorrências como as abaixo, em que não se observa valor de futuridade:

- (01) As professoras têm dificuldade em estar trabalhando esses conceitos.
- (02) A gente pode estar discutindo isso.

O Imperativo e o Subjuntivo, por sua vez, também parecem estar perdendo espaço para formas do Indicativo, especialmente no tempo Presente, de forma que ocorrências como as seguintes não causam mais estranhamento na língua falada (cf. SCHERRE et al, 2000; BORGES, 2004):

(03) Quer que eu **pago** pra você?

(04) **Faz** isso pra mim, por favor.

Já Barbosa (2008), que realizou um estudo diacrônico do Pretérito Perfeito Simples e Composto do Português Brasileiro do século XVI ao XX, verificou que o a forma composta foi se estabelecendo ao longo do tempo como um operador aspectual, deixando de expressar exclusivamente perfectividade e passando a expressar valores predominantemente iterativos e durativos.

Como podemos observar esses argumentos indicam que o sistema de tempos verbais do português está passando por uma reorganização, que ainda se configura como tendência, mas já pode ser entrevista, e que deve levar ao desaparecimento de certas formas, ao surgimento de outras, e a um deslocamento de funções. Assim, pesquisa torna-se relevante e se justifica devido uma série de questões que têm sido levantadas – mas ainda não resolvidas – a respeito do significado e dos usos das formas verbais, como algumas apontadas abaixo:

(i) quais os valores expressos pelas formas verbais do modo Indicativo no PM de Uberaba dos século XIX e XX?

(ii) existiriam dissimilaridades ou semelhanças em seus usos?

(iii) existem algumas formas que estão substituindo outras? Como, por exemplo, o Pretérito-Mais-Que-Perfeito Composto está realmente substituindo a forma simples correspondente como aponta Coan (2006) ou existem contextos que favorecem o emprego dessa forma? Isso acontece com outros tempos verbais como constatou, entre outros, Silva (1997), Scherre et al (2000); Borges (2004), Scherre (2007), Barbosa (2003, 2008)?

(iv) Os empregos das formas verbais do Indicativo apontados pelos gramáticos correspondem à realidade do sistema linguístico do PM? Ou as variações e rearranjos já provocaram um distanciamento entre o prescritivo e o descritivo, também nessa área?

(v) é possível encontrar no século XIX e XX vestígios da estrutura que originou as formas compostas, como apontou Barbosa (2008)?

Pelo que se acabou de apresentar, parece desejável um estudo do sistema temporal do Português Mineiro de Uberaba em textos jornalísticos a fim de que se possa estabelecer os contextos dos tempos e, a partir disso, chegar a uma descrição mais adequada sistema linguístico mineiro de Uberaba, que permita a comparação com outros sistemas, auxilie no estudo da história do PB e seja aplicável ao ensino/aprendizagem do português.

4.2. Revisão teórica das categorias verbais

Como embasamento teórico para os estudos das categorias verbais, partimos dos seguintes pressupostos para o desenvolvimento deste projeto: (i) as categorias modo, tempo e aspecto não estão apenas morfologicamente amalgamadas no português; relacionam-se estruturalmente também do ponto de vista semântico, ainda que muitas vezes seja possível identificar a predominância de uma delas na realização dos enunciados; (ii) as formas verbais têm significados básicos, sobre os quais podem operar adjuntos para dar origem aos valores que prevalecem nos diferentes contextos.

A primeira hipótese parece ser de aceitação consensual na bibliografia consultada. Das diferentes conceituações apresentadas, podemos extrair alguns pontos fundamentais:

(a) Para Ilari (1997), uma das dificuldades para o estudo semântico das formas verbais está na ausência de correspondência biunívoca entre os recursos expressivos e os conteúdos expressos, somada ao fato de que as construções que expressam tempo também exprimem modo e aspecto. Nem sempre é fácil separar os valores autenticamente temporais das expressões linguísticas de seus valores aspectuais e modais.

(b) Aspecto e tempo são categorias estreitamente relacionadas, mas que não se confundem, porque o aspecto não leva em conta o processo de enunciação, e o tempo, sim. Em outras palavras, somente o tempo é categoria dêitica; o aspecto leva em conta o intervalo em que se desenvolve o estado de coisas expresso pelo tempo.

(c) Quanto às categorias modo e tempo, afirma Fonseca (1977, p.110-111) que “longe de haver incompatibilidade, há, pelo contrário, íntima ligação entre o que era tradicionalmente designado por valores modais e valores temporais: a escolha de tempos

verbais tem uma incidência em toda a organização do discurso por parte do locutor e constitui, pois, um meio importante de modalização do enunciado. Só esta observação pode levar a uma efetiva compreensão dos chamados “empregos modais” de alguns “tempos”, como o futuro, o imperfeito, o condicional (uso modal, neste caso, já consagrado na própria designação da forma)”.

Com base nesses pontos, e para os fins deste projeto, adotaremos as seguintes definições:

- **Tempo** é uma categoria dêitica que expressa relações de anterioridade ou simultaneidade entre três momentos, o da fala, o do evento e o da referência.

- **Aspecto** é uma categoria não-dêitica, mediante a qual se quantifica o evento expresso pelo verbo ou se expressa a constituição interna de fases, momentos ou intervalos de tempo que se incluem nesse evento¹.

- **Modo** traduz uma apreciação qualitativa em relação ao enunciado, uma “tomada de posição do sujeito falante, ou a manifestação da vontade, sentimentos ou julgamento do sujeito gramatical” (LOBATO, 1971, p.99).

Se a ligação intrínseca entre as três categorias em geral não é questionada, a hipótese de valores semânticos não-ambíguos é fonte de divergências. Para autores como Lyons (1977) e Mira Mateus et al. (1989), as formas verbais são polissêmicas. O caso típico é o do Presente do Indicativo, que pode ser empregado para expressar futuro e na afirmação de fatos atemporais ou onitemporais. Apoiamo-nos em Ilari (1997) para avaliar duas alternativas de abordagem da questão da ambiguidade das formas verbais. Para o semanticista, postular tal ambiguidade seria problemático, pois, “à falta de adjuntos, os tempos verbais não são realmente polissêmicos, havendo interpretações privilegiadas, se não obrigatórias” (ILARI, 1997, p.25). A alternativa pressuposta em (ii) “encontra um obstáculo na abstração que seria exigida para se formularem os valores básicos” (p.25).

Outra dificuldade apontada por Ilari é que às vezes é o verbo que opera sobre o adjunto e não o contrário, como na frase (05), em que, segundo o autor, **agora** indica um momento passado:

(05) Agora, o paciente já não sentia dores, só um leve cansaço.

¹ Não vamos considerar aqui a distinção entre aspecto e modo de ser da ação. Para uma discussão a respeito, consulte-se Longo (1990, p. 102-104) e Barbosa (2008).

Excepcionalmente, as interpretações do adjunto e do verbo podem até ser não-equivalentes, como na frase **Eu sou você amanhã**, de um antigo comercial de uísque.

Em vista disso, concluímos que a interpretação semântica do sintagma verbal nem sempre está subordinada à interação com adjuntos, e preferimos não adotar a hipótese da polissemia, conscientes, portanto, de que teremos de buscar de representações mais abstratas.

Para interpretarmos semanticamente os tempos verbais do Português, baseamos-nos nos estudos de Corôa (1985) que procurou atribuir a cada tempo verbal do português uma definição única e não-ambígua. Fundamentada nos estudos de Reichenbach (1980), suas representações não levam em conta a interação verbo/adjunto temporal, nem o uso de auxiliares na expressão do tempo, centrando-se na interpretação fornecida pelo morfema modo-temporal do verbo. Partindo das possibilidades combinatórias dos três pontos temporais reichenbachianos - que ela denomina momentos do evento (ME), da fala (MF) e da referência (MR) - tenta verificar quais delas estão gramaticalizadas em nossa língua.

Desse reconhecimento da não concretude dos momentos resulta uma grande riqueza de possibilidades para o aproveitamento do esquema de Corôa na explicação dos diferentes usos e valores dos tempos verbais. Aceitando-se a não concretude dos momentos, elimina-se a necessidade de derivar contextualmente a sua interpretação.

4.3. Etapas e Estágio atual da pesquisa

Inicialmente realizamos uma revisão bibliográfica baseada em trabalhos que envolvem os tempos verbais do português e questões relacionadas à construção de corpus e a gêneros textuais. Simultaneamente estamos fazendo a organização do banco de dados, digitalizando os jornais antigos. Buscamos também fazer um levantamento dos gêneros que circulam nos jornais de cada época, suas características funcionais e sociais.

Por fim, a última etapa, ainda em andamento, é a análise linguística. Essa análise dos dados (dos tempos verbais) segue a metodologia variacionista: (a) o levantamento de uma amostra representativa dos tempos verbais estudados nos gênero jornalísticos; (b) análise de acordo com um grupo de fatores definidos a partir de hipóteses; (c) a quantificação dos dados analisados por meio de programas estatísticos (GOLDVARB). Por fim, será realizada a interpretação qualitativa dos resultados de acordo com os pressupostos teóricos que embasam este projeto.

Já pudemos observar variações semânticas na expressão da futuridade (exemplo 07, abaixo) e do “passado do passado” (exemplo 08, em que o Perfeito é utilizado no lugar do Mais que Perfeito):

(07) *Passa amanhã mais um aniversário da tomada da Bastilha. Data de comemoração universal...* (Gazeta / grifo nosso)

(08) *O telegramma chegou sexta-feira, porem o prefeito ja mandou sua decisão no dia anterior.* (Lavoura e Comercio / grifo nosso)

Portanto, como resultado deste trabalho, buscamos uma caracterização que parta de uma definição semântica abstrata para chegar a uma explicação dos possíveis usos e funções atribuídos às formas verbais do modo Indicativo no Português Mineiro de Uberaba.

5. A configuração pronominal dos demonstrativos a partir de jornais uberlandenses do século XX

Partindo de uma perspectiva sócio-discursiva dos fenômenos da linguagem, que enxerga a língua, portanto, como uma realidade dinâmica e multifacetada, resultante “das peculiaridades das experiências históricas e socioculturais do grupo que a usa: como ele se constitui, como é sua posição na estrutura socioeconômica, como ele se organiza socialmente, quais seus valores e visão de mundo, quais suas possibilidades de acesso à escola, aos meios de informação, e assim por diante” (FARACO, 1991, p.18), neste artigo serão apresentadas as hipóteses e os objetivos principais da pesquisa vinculada ao projeto trienal “Diversidade linguística em Minas Gerais: constituição de *corpus* jornalístico da cidade de Uberlândia do século XX”, desenvolvido por mim no Instituto de Letras e Linguística da Universidade Federal de Uberlândia, desde 2011.

Este projeto tem por finalidade, em linhas gerais: 1º) a construção e organização de um *corpus* digital referente aos exemplares da década de 1940 até o ano de 2000 do jornal “Correio de Uberlândia”, sendo que este material poderá auxiliar em pesquisas de caráter sociolinguístico, bem como em diversas outras áreas das Ciências Humanas; 2º) contribuir para o resgate do sócio-história de Uberlândia - cidade que é destaque no Triângulo Mineiro - através dos registros escritos expressos pelos diferentes gêneros textuais que compõem o jornal; 3º) Contribuir para a preservação do patrimônio histórico e

cultural da cidade de Uberlândia por meio da digitalização dos exemplares disponíveis no Arquivo Público de Uberlândia do jornal “Correio de Uberlândia”, haja vista tal digitalização ser uma importante maneira de restauração das informações de cunho social, político e cultural de determinadas épocas.

Atualmente, utilizando como *corpus* a seção de esportes de alguns exemplares do jornal “Correio de Uberlândia” da década de 1950, uma pesquisa linguística envolvendo o comportamento do sistema pronominal dos demonstrativos *este*, *esse* e *aquele*, foi iniciada. Tendo em vista que tal pesquisa é o foco da presente seção deste artigo, é importante que façamos uma breve revisão acerca destes pronomes para que se tornem mais claros os objetivos e as hipóteses que norteiam nosso estudo, ainda em fase preliminar.

5.1. Revisão bibliográfica dos pronomes demonstrativos

Os pronomes demonstrativos - de modo geral – indicam a posição dos seres em relação às três pessoas do discurso (“eu”, “tu/ você”, “ele”). Assim, temos, segundo descrições normativas², uma configuração ternária para o sistema pronominal dos demonstrativos na Língua Portuguesa, em que:

- formas de 1ª pessoa (*este/isto*) marcam referência ao campo do falante;
- formas de 2ª pessoa (*esse/isso*) marcam referência ao campo do ouvinte;
- formas de 3ª pessoa (*aquele/aquilo*) fazem referência ao que está fora do campo falante-ouvinte.

Quanto às estratégias de referenciação estabelecidas por tais pronomes, grosso modo, de acordo com Cunha (1986), os demonstrativos estabelecem referências dêiticas (do grego *deiktikós* = próprio para demonstrar; capacidade de mostrar um objeto sem nomeá-lo) e anafóricas (do grego *anaphorikós* = que faz lembrar, que traz à memória; capacidade de lembrar ao ouvinte ou ao leitor o que já foi dito ou o que irá se dizer).

Assim, as formas **este**, **esta**, **isto** indicam a) o que está perto da pessoa que fala e b) o tempo presente em relação à pessoa que fala. Além disso, são usadas para chamar a atenção sobre aquilo que dissemos ou vamos dizer. Já as formas **esse**, **essa**, **isso** indicam c) o que está perto da pessoa a quem se fala e d) o tempo passado ou futuro com relação à época em que se coloca a pessoa que fala. Cabe observar que são usadas

² Gramáticas Normativas do Português que foram pesquisadas: Cunha (1986), Infante (2002), Cipro Neto e Infante (2002) e Bechara (2006).

também para aludir ao que mencionamos ou para referirmos ao que foi dito por nosso interlocutor. Por fim, os pronomes **aquele**, **aquela**, **aquilo** denotam e) o que está afastado tanto da pessoa que fala como da pessoa a quem se fala e f) um afastamento no tempo de modo vago, ou uma época remota. Também podemos usar “aquele” para o que foi referido em primeiro lugar e, “este” para o que foi nomeado por último. Exemplo: **Caio** e **Mariana** foram viajar. Esta foi para Bahia e **aquele** para o Rio de Janeiro.

Dentre os gramáticos ora considerados, Cunha (1986) e Bechara (2006), embora apresentem uma distribuição ternária do sistema pronominal dos demonstrativos, são os únicos que reconhecem que nem sempre o uso desses pronomes corresponde à norma gramatical. Segundo Bechara (2006, p.167), no uso dos pronomes demonstrativos “muitas vezes interferem situações especiais que escapam à disciplina da gramática”.

5.2. O que dizem as pesquisas linguísticas quanto à configuração do sistema dos demonstrativos?

Pesquisas linguísticas apontam para uma simplificação do sistema dos demonstrativos, que de ternário (este/ esse/ aquele) teria passado a binário (este-esse/ aquele) no PB, devido ao apagamento da distinção entre o campo referencial do falante e do ouvinte, a fim de se criar apenas uma distinção dicotômica entre perto e longe (cf. Câmara Júnior, 1970, 1975; Castilho, 1978, 1993; Pavani, 1987; Roncarati, 2003; Marine, 2004; Pereira, 2005, Marine, 2009). Assim, as pesquisas apontam para um sistema binário no PB marcado, predominantemente, por “esse vs aquele”.

Cabe ressaltar que embora as culturas que operam com a escrita desenvolvam um padrão de língua codificado em gramáticas, cultivado por grande parcela dos letrados e ensinado pelas escolas, adquirindo, assim, um estatuto de estabilidade e permanência maior do que as outras variedades da língua, isso tudo se constitui apenas como um “refreador temporário de mudanças” (cf. FARACO, 2005, p.15), mas não como um inibidor, uma vez que a variação e a mudança são inerentes à língua.

5.3. Por que estudar o comportamento do sistema pronominal dos demonstrativos no Português Mineiro?

Nenhuma das pesquisas linguísticas sobre o sistema pronominal dos demonstrativos nos Português considerou dados do PM (Português Mineiro). Em geral, tais pesquisas foram realizadas com dados de São Paulo, Rio de Janeiro e de algumas

regiões do Nordeste. Realizar tal investigação é, portanto, contribuir não apenas para a descrição do Português Mineiro, especialmente, da região do Triângulo, mas, e sobretudo, para a descrição do PB (Português Brasileiro);

Nossa hipótese diante de tal estudo é que o sistema binário vai se confirmar tal como as demais pesquisas, apontando maior incidência da forma “esse”, ao menos no uso anafórico.

Para checarmos nossa hipótese, consideraremos os seguintes fatores para análise linguística: a) Função adjetiva ou substantiva do pronome; b) Uso dêitico subdivido em quatro tipos: dêixis temporal, dêixis espacial, dêixis textual e dêixis de memória³; c) Uso endofórico (anafórico), subdivido em cinco tipos: anáfora fiel, anáfora infiel, anáfora por nomeação, anáfora por elipse e anáfora de memória⁴ e d) Matização nas referências, afinal, muitas vezes falante/escritor, consciente ou inconscientemente, pode atribuir valores ao que deseja referenciar, ou seja, pode “matizar” seu “objeto de discurso”, ora mais, ora menos, positiva ou negativamente.

5.4. Estágio atual da pesquisa e considerações finais

Com o desenvolvimento deste projeto, já digitalizamos trinta anos de jornal – 1940 a 1960. Para a presente pesquisa, estamos separando todas as seções “Correio Esportivo” da década de 1950 (pelo que temos observado há uma média de 105 textos por ano). Partindo destes textos, estamos fazendo o levantamento das ocorrências das formas *este*, *esse* e *aquela* nos textos do “Correio Esportivo” para que, assim, possamos compor uma amostra significativa de dados e submetê-los às análises linguísticas, seguindo o modelo metodológico da sociolinguística laboviana.

Esperamos que ao término desta pesquisa, possamos descrever o sistema dos pronomes demonstrativos, tal como ocorriam no PM da década de 1950 (ternário ou binário?) e, por extensão, contribuir para a descrição do PB.

³ A dêixis de memória: proposta por Apothéloz (1995), ocorre quando um sintagma nominal pode referir-se *in absentia*, ou seja, na ausência de qualquer designação antecedente de seu referente e sem que este esteja presente na situação enunciativa; caracteriza-se também por remeter o interlocutor a um momento, situação ou acontecimento do passado. O locutor/escritor “dá pistas” do referente que será retomado pelo ouvinte/leitor, a fim de, ao narrar (contar) algo ao seu interlocutor, consiga remetê-lo espacial e temporalmente a um “evento” vivido pelo locutor/escritor.

⁴ A anáfora de memória é aquela que retoma o antecedente *in absentia*, ou seja, na ausência de qualquer designação explícita e prévia de seu referente. Este “resgate” pelo escritor/falante, trata-se de um assunto ou de uma idéia que deve ser, necessariamente, compartilhado pelo seu leitor/ouvinte.

6. A realização do objeto direto anafórico no português: um estudo em corpora constituídos a partir de programas de auditório

Sustentado pelos pressupostos teórico-metodológicos da Sociolinguística Variacionista, este estudo propôs estabelecer uma análise descritivo-analítico-comparativa da realização do **objeto direto anafórico** (cf. 09 e 10) no português brasileiro (PB) e no europeu (PE), enfocando, principalmente, o uso de pronomes pessoais com a respectiva função e o estatuto do objeto nulo.

(09) Mas o negócio do Célio que eu tava te contando, o negócio do sotaque, ele foi homenageado em Itu na volta né e o prefeito se reuniu com o pessoal pra homenageá-lo (PB.H.07.03)

(10) É a mesma coisa que... essa mania de deitar abaixo monumentos de pessoas que não... que não... não merecem o nosso respeito ou consideração, eu acho isso importante, esses monumentos e fotografias de homens de estado que foram maus, mas é bom que estejam num sítio onde nós possamos olhar \emptyset e dizer assim: olha o que este fez, mas ninguém deve fazer. Acho que é importante. O apagar a História é um erro crasso. (PE.F.06.09)

Objetivando verificar, no tocante ao fenômeno estudado, em que grau as referidas variedades linguísticas se distanciam e se esse distanciamento, caso haja, reflete apenas em uma distinção quantitativa entre as variedades ou se também causa algum reflexo quando visto de uma perspectiva qualitativa, assumiu-se a hipótese norteadora de que as distinções entre **PB** e **PE** são observadas, sobretudo, em uma análise quantitativa; em uma análise qualitativa, muito provavelmente, essas diferenças não se acentuarão.

Considerando que se trata de um estudo pautado no modelo teórico-metodológico da Sociolinguística Variacionista, proposto por Weinreich, Labov e Herzog (1968), Labov (1972, 1982, 1994, 2001), fez-se necessária a organização de *corpora* representativos das variedades linguísticas em estudo que permitissem, principalmente, o estabelecimento da comparação entre tais variedades. Por essa razão, a constituição dos *corpora* deve tomar com princípio o fato de que os falantes estarão sujeitos a situações semelhantes de produção de fala, uma vez que a organização de cada amostra deve seguir os mesmos critérios (ou pelo menos critérios semelhantes).

Dada a dificuldade de se encontrar para as variedades em estudo amostras linguísticas que tenham sido organizadas de modo que permitissem estabelecer uma comparação fiável, surgiu, então, a necessidade da organização dos próprios *corpora*. Tal organização se deu, então, a partir de programas de auditório⁵ – mais especificamente entrevistas ocorridas nesses programas – veiculados por emissoras de televisão de canal aberto e de circulação nacional nos países usuários das variedades acima referidas. Tal escolha se justifica pelo fato de esse gênero apresentar, em sua estrutura, uma rica variedade de perfis de falantes: informantes de ambos os sexos/gêneros, de idades diferentes e de diferentes níveis sócio-econômicos, possibilitando, dessa maneira, uma melhor observação dos níveis de variação.

Ressalta-se, ainda, o caráter popular dos *programas de auditório*, constituindo-se, na programação das emissoras de televisão, em um formato que permite a entrada do público nos estúdios ou nos locais preparados para gravação (quando não são transmitidos ao vivo), sendo, ainda, o público frequentemente convidado a participar do programa (SOUZA, 2004, p. 93).

Outro fato que merece destaque é a relação estabelecida entre o apresentador e seu(s) convidado(s): o apresentador, sendo responsável por garantir a descontração de seus convidados, assim como por fazer com que os membros da plateia se sintam o mais confortável possível, assume papel fundamental. Os convidados são recebidos em um ambiente que, embora descontraído, apresenta sofisticação e se mostra propício para que possam estabelecer um diálogo em que, seguindo o pensamento de Rosário (2008), “intimidades são reveladas e banalidades multiplicadas”.

O trabalho com o gênero denominado *programas de auditório* permitirá, ainda, verificar o uso linguístico, seguindo os passos de Duarte (1989, p. 20), em um registro de fala que atinge os países de ponta a ponta, exercendo sobre a comunidade linguística, simultaneamente, uma força inovadora e normalizadora, possibilitando, conforme proposto por Labov (1972, p. 63), a organização de *corpora* tomando como ponto de partida a língua em seu contexto social.

Nessa mesma linha pensamento, Ávila (1999), referindo-se não só à televisão, mas também ao rádio, afirma que se constituem em meios de comunicação de massa que contribuem para a difusão de uma língua, da mesma forma que para sua consolidação como língua nacional, contribuindo para a normalização das línguas que transmitem.

⁵ Da programação da televisão brasileira, foram utilizados os programas Domingão do Faustão e Programa do Jô; da televisão portuguesa, foram utilizados os programas Herman 2010, Companhia das Manhãs e Vida Nova.

Feitas essas breves explanações sobre o fenômeno estudado, assim como sobre a organização dos *corpora*, são apresentados os resultados gerais da pesquisa.

6.1. Objeto direto no PB e no PE: alguns resultados

Organizados os *corpora* da investigação, o trabalho desenvolvido por Arruda (2012) controlou a natureza do antecedente do **OD**: de um lado, antecedente **SN (OD(SN))** e, de outro, antecedente **oracional (OD(or))** (ou, como denominado por alguns linguistas, **sentencial**), sendo consideradas as seguintes formas variantes do fenômeno em estudo: **OD nulo, clítico, pronome lexical, SN pleno, SN com determinante modificado e demonstrativo**. Feitos, então, o levantamento dos dados e sua codificação de acordo com os grupos de fatores considerados – **animacidade e especificidade** do antecedente, **grau de transitividade da sentença, estrutura sintática** em que é projetado o **OD**, **topicalização (ou não) do antecedente**, e **sexo/gênero do informante** (apenas os três últimos foram considerados em casos de **OD(or)**) – os dados foram submetidos à análise estatística, tendo frequências, assim como resultados de peso relativo, obtidos a partir do programa GoldVarb X⁶.

A tabela apresentada a seguir possibilita visualizar os resultados gerais da distribuição das variantes do **OD(SN)** no **PB** e no **PE**.

Tabela 2 – Resultados gerais para a realização das variantes de OD(SN) no PB e no PE

Variantes	Variedades			
	PB		PE	
	Quant.	%	Quant.	%
Nulo	146	64,5	163	57
Clítico	07	3	71	25
Pron. Lexical	15	6,5	02	0,7
SN Pleno	39	17	24	8,5
SN c/ Det. Modificado	13	6	12	4,3
Demonstrativo	07	3	13	4,5
TOTAL	227	100	285	100

Ainda que a tabela apresente a frequência obtida para todas as variantes consideradas no estudo de Arruda (2012), a discussão, aqui, estará restrita à frequência das variantes **nulo, clítico e pronome lexical**.

Nesse sentido, os dados apresentados na tabela acima permitem verificar que, tanto no **PB** quanto no **PE**, o objeto nulo (cf. (11) e (12)) é a forma preferida pelos falantes

⁶ Aqui serão apresentados apenas os resultados gerais do trabalho de Arruda (2012).

para a realização do **OD** anafórico, tendo um SN como antecedente. O resultado verificado para essa variante – ainda que com um percentual um pouco superior no **PB** – supera, tanto em uma quanto em outra variedade, a soma das demais variantes consideradas no estudo.

(11) mas em contra partida sou muito mais criterioso na forma como gasto meu tempo [...] mas saboreio \emptyset bem melhor (PE.M.04.02)

(12) esse é o auto bronzeador, é aquele que, especialmente na Europa, italiano é doido por isso, ele usa \emptyset no inverno, no verão ele também usa e não precisa você ir, se esticar no sol não. (PB.M.01.03)

Em relação ao **PB**, esses índices convergem com os identificados por outros estudos, quer orientados pela perspectiva variacionista – como os de Duarte (1986 e 1989), Matos (2005) e Arruda (2006) –, quer pela perspectiva gerativista – como o de Cyrino (1999) –, ou ainda pela perspectiva da Sociolinguística Paramétrica – como o de Freire (2000). Esses estudos também apontam para a preferência dos falantes pela variante objeto nulo como forma de realização do **OD** em posição de anáfora. No entanto, chama-nos a atenção a alta frequência verificada para essa variante no **PE**.

Passando agora o foco às variantes **clítico** e **pronome lexical**, os índices verificados para as variedades do português não apresentam igual semelhança.

No tocante à realização do **OD** anafórico por uma forma pronominal, duas variantes foram consideradas: **clítico** (cf. (13) e (14)) e **pronome lexical** (cf. (15)). No que diz respeito à primeira variante (forma prescrita pela Gramática Tradicional, portanto tida como de maior prestígio linguístico), os índices identificados se distanciam bastante quando comparadas as duas variedades do português: 3% no **PB** e 25% no **PE**. Em relação ao **PB**, tal frequência já era esperada, uma vez que se assemelha aos índices verificados por Arruda (2006) e confirma a afirmação de Duarte (1986 e 1989) de que o índice de uso de clíticos na posição de objeto direto, em se tratando do uso feito por falantes com maior grau de escolaridade, seria de aproximadamente **5%**. No entanto, ainda que bastante superior à frequência do **PB**, não se esperava, para o **PE**, uma frequência inferior à identificada para a variante **OD nulo**.

(13) Uma mulher com uma vida comum e foi o que eu passei pro Veríssimo e ele então inventou esses absurdos todos do Diabo, dessas pernas que tem vida própria, uma pessoa que ela carrega, aliás as pernas **a** carregam né? (PB.F.06.03)

(14) há uma pleiade enorme de músicas que... de estilos musicais que eu aprecio e que me dão muita felicidade, muito prazer em ouvi-**los**. (PE.M.07.03)

(15) se ela [a irlandesa], chama super legítimo, se ela chama a polícia na minha casa eu também vou tomar medidas pra incomodar **ela** na medida em que ela tá me incomodando (PB.F.12.03)

Em relação ao uso do **pronome lexical** como manifestação do **OD anafórico** no **PB**, ainda que aponte para uma frequência baixa (6,5%), é possível verificar que esse índice supera o dobro do identificado quando se trata de um **clítico**. Em relação ao **PE**, ainda que os dados apontem para uma frequência inferior a 1% (apenas duas ocorrências), essa frequência pode ser, de fato, considerada **zero**, dados os casos em que se dão. Ambas se dão em contexto em que se tem, nos dizeres de Duarte (1986 e 1989), uma estrutura complexa, ou seja: verbo + objeto direto + infinitivo verbal (oração infinitiva) (cf. (16) e (17)), em que a forma pronominal de terceira pessoa funciona como sujeito de infinitivo.

(16) Eu acho que de... é de gostar e de... eu fico doido... agora, com o Youtube, eu já vi tudo o que há de**la**. Já vi **ela** a cantar em TelAviv (PE.F.09.06)

(17) a Amália, como não se ouvia, cá está a história, nunca se encheu dela, nunca achou que ela era extraordinária, mais tarde tinha descoberto o... casualmente ouviu **ela** cantar uma canção francesa (12.12)

Ainda que tenham sido aqui apresentados apenas os resultados gerais, ao estabelecer a análise a partir do cruzamento entre os grupos de fatores, o estudo de Arruda (2012) pode verificar, sobretudo em relação ao **OD nulo** que as duas variedades linguísticas seguem na mesma direção, confirmando, assim, a hipótese anteriormente apresentada.

7. Considerações Finais

Como podemos observar, ao longo deste artigo, o que une os resultados apresentados é o conhecimento e a reflexão sobre diferentes caminhos metodológicos e teóricos de pesquisas linguísticas em diferentes áreas que culminam na montagem de corpora e bancos de dados sobre o Português, seja na sua perspectiva histórica, com estudos que contemplam desde o português arcaico, passando por outros períodos da língua, como o português do século XIX, por exemplo, até estudos sobre o português atual, tanto na perspectiva sincrônica, quanto na perspectiva diacrônica, comparando-se diferentes períodos e variedades dessa língua.

8. REFERÊNCIAS

- ARRUDA, N. C. A realização do objeto direto anafórico em línguas românicas: um estudo sincrônico no português e no espanhol. 2012. 150f. Tese (Doutorado em Linguística e Língua Portuguesa). Faculdade de Ciências e Letras, UNESP, Araraquara, 2012.
- ARRUDA, N. C. A realização do objeto direto no português brasileiro culto falado: um estudo sincrônico. 2006. 201f. Dissertação (Mestrado em Linguística e Língua Portuguesa). Faculdade de Ciências e Letras, UNESP, Araraquara, 2006.
- ÁVILA, R. Radio, televisión y lengua: hacia la norma internacional del español. In: MORALES, A.; CARDONA, Julia; LOPES MORALES, H.; FORASTIERI, E. **Estudios de lingüística hispánica: homenaje a María Vaquero**. Puerto Rico: Editorial de la Universidad de Puerto Rico, 1999, p. 73-89.
- BARBOSA, J. B. **Tenho feito/fiz a tese: uma proposta de caracterização do Pretérito Perfeito no português**. 2008, 272 p. Tese (Doutorado em Linguística e Língua Portuguesa). Faculdade de Ciências e Letras, UNESP, Araraquara, 2008.
- BARBOSA, J. B. ; LONGO, B. N. O. As formas futurizadas no português contemporâneo. In: XIV Encontro do CELLIP.XVI, 2001, **Anais do XIV Encontro do CELLIP**.CD-ROOM.
- BARBOSA, J. B. **Os tempos do pretérito no português brasileiro: perfeito simples e perfeito composto**. 2003. 115f. Dissertação (Mestrado em Linguística e Língua Portuguesa). Faculdade de Ciências e Letras, UNESP, Araraquara, 2003.
- BECHARA, E. **Moderna gramática portuguesa**. 37.ed. Rio de Janeiro: Lucerna, 2006. (edição revista e ampliada).
- BERLINCK, R. de A.; BARBOSA, J.B.; MARINE, T. de C. Reflexões teórico-metodológicas sobre fontes para o estudo histórico da língua. **Revista da ABRALIN**, v. 7, 2008, p. 53-79.
- BORGES, P. Formas imperativas em tiras de jornais paulistas. 2004. Dissertação (Mestrado em Linguística e Língua Portuguesa). Faculdade de Ciências e Letras, UNESP, Araraquara, 2004.
- CÂMARA JR., J.M. **Estrutura da Língua Portuguesa**. Petrópolis: Vozes, 1970.
- CÂMARA JR., J.M.. **História e Estrutura da Língua Portuguesa**. Rio de Janeiro: Padrão, 1975.

- CASTILHO, A. T. de. **Análise preliminar dos demonstrativos na norma culta de São Paulo**. Anais de Seminários do GEL, 1978, p.30-35.
- CASTILHO, A. T. de. Os mostrativos do português falado. In: **Gramática do Português Falado**. Campinas: As abordagens, v. 3, 1993, p.119-145.
- CIPRO NETO, P; INFANTE, U. **Gramática da Língua Portuguesa**. São Paulo: Scipione, 2002.
- CUNHA, C. **Gramática da língua portuguesa**. Rio de Janeiro: Feriame, 1986.
- CYRINO, S.M.L. **O objeto nulo no português do Brasil: um estudo sintático-diacrônico**. Londrina: UEL, 1997.
- CYRINO, S.M.L. Elementos nulos pós-verbais no português brasileiro oral contemporâneo. In: NEVES, M. H. M. **Gramática do português falado: novos estudos**. Campinas: UNICAMP, 1999, v. VII.
- DUARTE, M. E. L. **A perda do princípio ´evite pronome` no português brasileiro**. 1995. Tese (Doutorado em Linguística). IEL, UNICAMP, Campinas , 1995.
- DUARTE, M. E. L. Clítico acusativo, pronome lexical e categoria vazia no português do Brasil. In: TARALLO, F. **Fotografias sociolingüísticas**. Campinas: Pontes, 1989.
- DUARTE, M. E. L. **Varição e sintaxe: clítico acusativo, pronome lexical e categoria vazia no português do Brasil**. 1986. 73f. Dissertação (Mestrado em Linguística Aplicado ao Ensino de Línguas). PUC, São Paulo, 1986.
- FARACO, C. A. **Linguística histórica**. São Paulo: Ática, 1991 (2005).
- FREIRE, G. C. Os clíticos de terceira pessoa e as estratégias para sua substituição na fala culta brasileira e lusitana. 2000. 107f. Dissertação (Mestrado em Letras Vernáculas (Língua Portuguesa). Faculdade de Letras, UFRJ, Rio de Janeiro, 2000.
- ILARI, R. **A expressão do tempo em português**. São Paulo, Contexto, 1997.
- INFANTE, U. **Curso de Gramática Aplicada aos Textos**. São Paulo: Scipione, 2002.
- LABOV, W. Building on empirical foundations. In: LEHMANN, W.; MALKIEL, Y. (org.). **Perspectives on historical linguistics**. Philadelphia: John Benjamins Publishing Company, 1982.
- LABOV, W. **Padrões sociolingüísticos**. Trad.: Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues Cardoso. São Paulo: Parábola, 2008.
- LABOV, W. **Principles of linguistic change**, Vol. 2: Social factors. Malden, MA: Blackwell, 2001.
- LABOV, W. **Principles of linguistic change**, Vol.1: internal factors. Oxford: Blackwell, 1994.
- LABOV, W. **Sociolinguistic Patterns**. Philadelphia : University of Pennsylvania Press, 1972.
- LOBATO, M. L. Os verbos auxiliares em Português contemporâneo. **Análises linguísticas**. Petrópolis: Vozes, 1975, vol LXV, p.27-91.
- LONGO, B. N. O. **A auxiliaridade e a expressão do tempo em português**. Doutorado. Araraquara, Unesp.1990.

MARINE, T. C. **O binarismo dos pronomes demonstrativos no século XX: este vs. aquele ou esse vs. aquele?**, 2004. 138 f. Dissertação (Mestrado em Lingüística e Língua portuguesa) – Faculdade de Ciências e Letras, UNESP, Araraquara, 2004.

MARINE, T. de C. **Um estudo sócio-discursivo do sistema pronominal dos demonstrativos no Português Contemporâneo**. 2009. 219f. Tese (Doutorado em Linguística e Língua Portuguesa) - Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, 2009.

MATOS, M. Z. M. S. A expressão do objeto direto anafórico nos falares urbanos itabienses. 2003. 171f. Dissertação (Mestrado em Linguística e Língua Portuguesa). Faculdade de Ciências e Letras, UNESP, Araraquara, 2005.

PAVANI, S. **Os demonstrativos este, esse e aquele no português culto de São Paulo**, 1987. Dissertação (Mestrado em Lingüística). Instituto de Estudos da Linguagem, Universidade de Campinas, Campinas, 1986, 90 p.

PEREIRA, H. B. **“Esse” versus este no Português Brasileiro e Europeu**. 2005. 110f. (Mestrado em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Ciências e Letras, USP, São Paulo, 2005.

RONCARATI, C. Os mostrativos na variedade carioca falada. In: PAIVA, M. da C.; DUARTE, M. E. L. (Org.) **Mudança lingüística em tempo real**. Rio de Janeiro: Contra Capa Liv., 2003.

ROSÁRIO, N. M. do. Do *talk show* ao televisivo: mais espetáculo, menos informação. In: **Em questão**. Porto Alegre, 2008, v. 14, n. 2, p. 149-162, jul./dez.

SCHERRE, M. M. P. Aspectos sincrônicos e diacrônicos do imperativo gramatical no português brasileiro. Alfa (ILCSE/UNESP), v. 51, p. 189-222, 2007.

SCHERRE, M. M. P. et al. Restrições sintáticas e fonológicas na expressão variável do imperativo no português do Brasil. **Boletim da Associação Brasileira de Lingüística (ABRALIN)**, Florianópolis, p. 1333-1347, 2000.

SILVA, A. F. **A expressão da futuridade na língua falada**. 1997. Tese (Doutorado em Linguística). Campinas, UNICAMP, 1997.

SOUZA, J. C. A. **Gêneros e formatos na televisão brasileira**. São Paulo: Summus, 2004.

WEINREICH, U.; LABOV, W.; HERZOG, M. **Empirical foundations for a theory of language change**. In LEHMAN, W.P.e MALKIEL, Y., (Ed.) Directions for historical linguistics. Austin: University of Texas Press, 1968.