



Universidade Federal da Bahia
Instituto de Letras
Programa de Pós-Graduação em Letras e Lingüística

Rua Barão de Geremoabo, nº147 - CEP: 40170-290 - Campus Universitário Ondina Salvador-BA
Tel.: (71) 3263 - 6256 – Site: <http://www.ppgll.ufba.br> - E-mail: pgletba@ufba.br



**CONCORDÂNCIA VERBAL EM PORTUGUÊS:
O QUE NOS REVELA O PERÍODO ARCAICO?**

por

PEDRO DANIEL DOS SANTOS SOUZA

Orientadora: Prof^a. Dr^a. Rosa Virgínia Mattos e Silva

Salvador
2005



Universidade Federal da Bahia
Instituto de Letras
Programa de Pós-Graduação em Letras e Lingüística

Rua Barão de Geremoabo, nº147 - CEP: 40170-290 - Campus Universitário Ondina Salvador-BA
Tel.: (71) 3263 - 6256 – Site: <http://www.ppgll.ufba.br> - E-mail: pgletba@ufba.br



**CONCORDÂNCIA VERBAL EM PORTUGUÊS:
O QUE NOS REVELA O PERÍODO ARCAICO?**

por

PEDRO DANIEL DOS SANTOS SOUZA

Orientadora: Prof^a. Dr^a. Rosa Virgínia Mattos e Silva

Dissertação apresentada ao Programa de Pós-Graduação em Letras e Lingüística do Instituto de Letras da Universidade Federal da Bahia, como requisito parcial para a obtenção do título de Mestre em Letras.

Área de concentração: Lingüística Histórica.

Salvador
2005

Biblioteca Central Reitor Macedo Costa – UFBA

S729 Souza, Pedro Daniel dos Santos.

Concordância verbal em português: o que nos revela o período arcaico? / por Pedro Daniel dos Santos Souza. - 2005.
95 f.

Orientadora : Profª Drª Rosa Virgínia Mattos e Silva.
Dissertação (mestrado) - Universidade Federal da Bahia, Instituto de Letras,
2005.

1. Lingüística histórica. 2. Sociolingüística. 3. Língua portuguesa - Concordância. 4. Língua portuguesa – Português arcaico – Séc. XIII-XIV. 5. Língua portuguesa – Variação. 6. Língua portuguesa – Gramática. I. Silva, Rosa Virgínia Mattos e. II. Universidade Federal da Bahia. Instituto de Letras. III. Título.

CDU – 81-112
CDD – 410

PEDRO DANIEL DOS SANTOS SOUZA

**CONCORDÂNCIA VERBAL EM PORTUGUÊS:
O QUE NOS REVELA O PERÍODO ARCAICO?**

Dissertação apresentada ao Programa de Pós-Graduação em Letras e Linguística do Instituto de Letras da Universidade Federal da Bahia, como requisito parcial para a obtenção do título de Mestre em Letras.

Área de concentração: Linguística Histórica.

Aprovada em 06 de abril de 2005.

BANCA EXAMINADORA

Rosa Virgínia Barretto de Mattos Oliveira e Silva - Orientadora _____
Doutora em Linguística pela Universidade de São Paulo
Universidade Federal da Bahia

Maria Marta Pereira Scherre _____
Doutora em Linguística pela Universidade Federal do Rio de Janeiro
Universidade de Brasília

Tânia Conceição Freire Lobo _____
Doutora em Filologia e Língua Portuguesa pela Universidade de São Paulo
Universidade Federal da Bahia

A minha mãe,
exemplo de compreensão.

AGRADECIMENTOS

À Prof^a. Dr^a. Rosa Virgínia Mattos e Silva, pela orientação sempre estimulante, despertando em mim o amor ao português arcaico.

Ao Prof. Dr. Américo Venâncio Lopes Machado Filho, pela satisfação em ajudar sempre que era solicitado, além do grande incentivo.

Aos colegas do Grupo de Pesquisa PROHPOR, pelo apoio, amizade e, sobretudo, a troca de experiência acadêmica.

Aos meus pais, Pedro e Jacira, pela formação e total dedicação à realização dos meus sonhos.

Aos meus estimados irmãos Jadilson, Jaciane, Gleidson, Daniela e Jacó, pelo incentivo constante e a confiança, além do grande carinho.

Os homens fazem a língua, e
não a língua os homens.

Fernão de Oliveira

RESUMO

Este trabalho investiga a variação da concordância verbal na primeira fase do período arcaico da língua portuguesa (séculos XIII e XIV), a partir de um *corpus* constituído por textos literários e não-literários (oficiais, particulares e institucionais) representativos da produção medieval portuguesa. Os dados, submetidos ao programa VARBRUL, confirmam a existência de uma regra sintática variável já na primeira fase do período arcaico, definida pela influência de fatores de ordem morfo-fônica, sintática e semântica, além de social. As variáveis tipo de verbo, saliência fônica, posição do sujeito em relação ao verbo, indicação do plural no sujeito, caracterização semântica do sujeito, tipo de texto, realização do sujeito e língua original do texto demonstram desempenhar um papel significativo no uso da concordância verbal. Com base nos resultados, verbo inacusativo, menor nível de saliência, posposição ao verbo, sujeito [- humano], texto literário e sujeito realizado tendem a desfavorecer a aplicação da regra de concordância verbal. Assim, a variação na primeira fase do período arcaico não é aleatória, mas condicionada por fatores que também têm sido considerados relevantes para a variação no português brasileiro.

Palavras-chave: Lingüística histórica; português arcaico; concordância verbal; variação; regra variável.

ABSTRACT

This work investigates the variation of verbal concordance in the first phase of the Old Portuguese, i.e. the 13th and 14th centuries, from a corpus composed by literary and non-literary texts (official, private and institutional texts) representing the medieval Portuguese production. The data submitted to the VARBRUL confirm the existence of a variable syntactical rule already used in the Old Portuguese, defined by morpho-phonetical, syntactical and semantic factors, besides the social ones. The variants type of verb, phonetics prominence, position of the subject in relation to the verb, indication of the plural form in the subject, semantical characterization of the subject, type of text, realization of the subject and original language of the text show an important role in the use of the verbal concordance. Based on the results, non-accusative verb, lesser level of prominence, post position to the verb, [-human] subject, literary text and realized subject have an unfavorable tendency to the application of the rule of the verbal concordance. Thus, the variation in the first phase of the Old Portuguese is not aleatory, but conditioned by factors which have also been considered relevant to the variation of the Brazilian-Portuguese.

Keywords: Historical Linguistics; Old Portuguese; verbal concordance, variation, variable rule.

LISTA DE TABELAS

Tabela 1 – Variável <i>tipo de verbo</i>	79
Tabela 2 – Cruzamento das variáveis <i>tipo de verbo e saliência fônica</i>	81
Tabela 3 – Variável <i>saliência fônica</i>	82
Tabela 4 – Variável <i>posição do sujeito em relação ao verbo</i>	84
Tabela 5 – Variável <i>posição do sujeito em relação ao verbo (dados amalgamados)</i>	85
Tabela 6 – Variável <i>indicação do plural no sujeito</i>	86
Tabela 7 – Variável <i>caracterização semântica do sujeito</i>	88
Tabela 8 – Variável <i>tipo de texto</i>	90
Tabela 9 – Variável <i>tipo de texto (dados amalgamados)</i>	91
Tabela 10 – Variável <i>realização do sujeito</i>	94
Tabela 11 – Variável <i>língua original do texto</i>	95

LISTA DE ABREVIATURAS E SIGLAS

DNAM	Documentos notariais editados por Ana Maria Martins
DNCM	Documentos notariais editados por Clarinda de Azevedo Maia
FRAX	Foro Real de Afonso X
FLOS	Flos Sanctorum
ORTO	Orto do Esposo
P3	Terceira pessoa (terceira pessoa do singular)
P6	Sexta pessoa (terceira pessoa do plural)
PB	Português brasileiro
PE	Português europeu
SN	Sintagma nominal
SPrep	Sintagma preposicionado
TASL	Testamento de Afonso II – Lisboa
TAST	Testamento de Afonso II – Toledo
V	Verbo
VARBRUL	<i>Variable rules</i>

SUMÁRIO

1	INTRODUÇÃO	12
2	DA VISÃO “PREÇEITIVA” À VISÃO DESCRITIVA	17
2.1	A TRADIÇÃO GRAMATICAL LUSO-BRASILEIRA	18
2.2	A VARIAÇÃO NO PERÍODO ARCAICO: “OUTROS OLHARES”	27
3	CONSIDERAÇÕES TEÓRICO-METODOLÓGICAS	43
3.1	TEORIA DA VARIAÇÃO E MUDANÇA LINGÜÍSTICA	43
3.2	DELIMITAÇÃO DO PROBLEMA	46
3.3	DO FALADO PARA O ESCRITO: A TRANSMISSÃO TEXTUAL	50
3.4	<i>CORPUS</i> : CARACTERIZAÇÃO	55
3.4.1	Testamento de D. Afonso II	56
3.4.2	Documentos notariais editados por Clarinda de Azevedo Maia	57
3.4.3	Documentos notariais editados por Ana Maria Martins	58
3.4.4	Foro Real de Afonso X	60
3.4.5	Orto do Esposo	62
3.4.6	Flos Sanctorum	63
3.5	VARIÁVEIS LINGÜÍSTICAS E EXTRALINGÜÍSTICAS	64
3.5.1	Realização do sujeito	64
3.5.2	Posição do sujeito em relação ao verbo	65
3.5.3	Caracterização semântica do sujeito	66
3.5.4	Saliência fônica	66

3.5.5	Tipo de verbo	68
3.5.6	Indicação do plural no sujeito	69
3.5.7	Língua original do texto	69
3.5.8	Tempo	70
3.5.9	Tipo de texto	71
3.6	SUPORTE QUANTITATIVO	72
4	ANÁLISE DO <i>CORPUS</i>	74
4.1	TIPO DE VERBO	79
4.2	SALIÊNCIA FÔNICA	82
4.3	POSIÇÃO DO SUJEITO EM RELAÇÃO AO VERBO	83
4.4	INDICAÇÃO DO PLURAL NO SUJEITO	86
4.5	CARACTERIZAÇÃO SEMÂNTICA DO SUJEITO	88
4.6	TIPO DE TEXTO	89
4.7	REALIZAÇÃO DO SUJEITO	93
4.8	LÍNGUA ORIGINAL DO TEXTO	95
5	CONSIDERAÇÕES FINAIS	97
	REFERÊNCIAS	101

1 INTRODUÇÃO

Há alguns anos, a variação na concordância verbal¹ em português tem sido objeto de análises exaustivas, seja em abordagens variacionistas, seja sob um enfoque gerativista (cf. GALVES, 2001). Independentemente do modelo teórico utilizado, as análises até agora realizadas priorizaram o aspecto sincrônico. No entanto para uma visão mais completa desse fenômeno lingüístico, faz-se necessário estudá-lo, exaustivamente, sob outra dimensão – a diacrônica. Isso não significa que a variação na concordância verbal nunca tenha sido estudada do ponto de vista histórico, mas sim que é preciso analisá-la observando um *corpus* representativo da documentação medieval portuguesa como sugere Mattos e Silva (1986), a fim de tornar mais preciso o conhecimento do português no caminho de sua história e, conseqüentemente, compreender as variáveis que controlam a variação do fenômeno lingüístico sob investigação.

No artigo *Idéias para a história do português brasileiro: fragmentos para uma composição posterior*, apresentado durante o “I Seminário para a História do Português Brasileiro”, Mattos e Silva (1998a) busca uma compreensão do português brasileiro (PB), externando o que supõe que sejam “algumas vias para entrever-se o passado desse português brasileiro” que deverão movimentar-se em direção à recuperação tanto histórico-social quanto lingüística do Brasil.

Apoiando-se na afirmação de que “as *histórias* que se escrevem são sempre *uma* história, reconstruída e escrita a partir dos condicionamentos teóricos,

¹ A expressão *concordância verbal*, assim como *concordância verbo-nominal*, é usada para fazer referência ao mecanismo sintático que estabelece a coesão sintática entre o sintagma nominal (SM) sujeito e o verbo (V) núcleo do predicado, já que ambos devem estar no mesmo número e referir-se à mesma pessoa gramatical, como prescreve o padrão normativo, embora este padrão também registre casos variáveis que são chamados de “facultativos”.

metodológicos, empíricos e ideológicos de seus autores” (MATTOS E SILVA, 1998a, p. 34), a autora considera que a reconstrução e subsequente escrita da história do PB passam pela necessidade de

reconstituir uma história social lingüística do Brasil: uma sociolingüística histórica (ou sócio-história lingüística) e uma história lingüística, ou seja, a história das mudanças lingüísticas que fizeram e fazem o português brasileiro apresentar as características que tem, o seu perfil próprio, a sua gramática. (1998a, p. 39)

Assim, visando a contribuir com “algumas idéias para a concretização de uma história do português brasileiro”, Mattos e Silva (1998a, p. 40) apresenta quatro grandes campos de pesquisa que, embora no seu processo de realização possam ser desenvolvidos de forma autônoma, necessariamente encontram-se interligados, a saber:

- a) *reconstrução de uma história social lingüística do Brasil*, partindo de fontes históricas múltiplas;
- b) *reconstrução de uma sócio-história lingüística ou de uma sociolingüística histórica*, que se moverá numa perspectiva de estabelecer correlações entre fatores lingüísticos e fatores sociais;
- c) *reconstrução diacrônica no interior das estruturas da língua portuguesa em direção ao português brasileiro*, buscando, através de uma *história interna*, descrever e explicitar ou explicar o encaixamento no interior das estruturas e a difusão da variante em mudança pela estrutura;
- d) *comparação entre o português brasileiro e o português europeu*, a fim de verificar se as mudanças ocorridas no PB já estariam prefiguradas ou encaixadas nele.

Uma grande conquista “desse programa de trabalho coletivo”, que deve ser desenvolvido em quatro campos de pesquisa, conforme explicitação acima, será a possibilidade de demonstrar se o que tipifica o PB decorre de *derivas* antigas ou resulta das peculiaridades sócio-históricas que condicionaram sua formação. Como conclui Mattos e Silva (1998a, p. 49), esse é um grande passo para “compreender um aspecto fundamental da nossa história pregressa”.

Diante dessas questões, vale ressaltar que o presente trabalho, visando a atender a proposta explicitada no campo *c*, investiga a questão diacrônica da concordância verbal no passado do português europeu, mais especificamente a primeira fase do período arcaico (séculos XIII e XIV), embora os resultados possam também trazer algumas contribuições quanto à discussão que envolve as hipóteses explicativas sobre as origens e formação do PB. Possivelmente, os dados poderão contribuir para a confirmação do problema do encaixamento histórico da variação da concordância, tópico que, segundo Mattos e Silva (1998a, p. 48), é “pedra de toque (...) da gramática do português brasileiro”.

Inicialmente, no capítulo DA VISÃO “PREÇEITIVA” À VISÃO DESCRITIVA, serão discutidas as questões postas pela tradição gramatical luso-brasileira, visando a verificar de que forma o fenômeno lingüístico em foco tem sido encarado por esse tipo de abordagem. Em seguida, serão apresentados alguns estudos que têm documentado a variação no período arcaico da língua portuguesa e as conclusões a que se podem chegar a partir das observações feitas.

O capítulo seguinte, CONSIDERAÇÕES TEÓRICO-METODOLÓGICAS, apresenta o aparato teórico no qual se fundamenta a pesquisa, ou seja, a “Teoria da variação e mudança” proposta no texto inaugural de Uriel Weinreich, William Labov e Marvin I. Herzog (1968), e sintetiza os princípios norteadores da análise aqui

realizada, além de procurar delimitar o problema investigado. Como se trata de um trabalho de natureza diacrônica, também serão feitas algumas reflexões sobre a transmissão textual a partir das relações entre o “falado” e o “escrito”, já que o *corpus* é constituído de textos representativos dos séculos XIII e XIV, portanto, primeira fase do período arcaico da língua portuguesa. Posteriormente, faz-se uma breve caracterização do *corpus*, destacando aspectos referentes à datação dos textos, língua em que foram escritos, conteúdo, dentre outros aspectos. Para encerrar essas considerações de ordem teórico-metodológica, as variáveis lingüísticas e extralingüísticas que orientaram a presente investigação serão elencadas, buscando, na medida do possível, justificar a motivação para a escolha dessas variáveis, da mesma forma em que se faz uma apresentação do suporte quantitativo utilizado para quantificar e analisar os dados, o VARBRUL.

A ANÁLISE DO CORPUS, tópico do capítulo posterior, fundamenta-se nos resultados obtidos a partir da exposição dos dados ao programa computacional escolhido para servir de suporte estatístico. Assim, serão discutidos os contextos que favorecem e/ou desfavorecem a concordância verbal na primeira fase do período arcaico, buscando chegar a conclusões plausíveis, não deixando de levar em consideração a natureza do *corpus* que fundamenta a presente investigação. Nesta seção, a discussão partirá das variáveis lingüísticas e extralingüísticas selecionadas pelo VARBRUL, apresentando os índices gerais de concordância, seguidos dos percentuais e, sobretudo, pesos relativos atribuídos a cada variável. Diante disso, as tabelas que apresentam os resultados obtidos com a quantificação das ocorrências tornam-se um instrumento valioso na compreensão do fenômeno lingüístico investigado.

Finalmente, as CONSIDERAÇÕES FINAIS, seguidas das REFERÊNCIAS consultadas que serviram de suporte para esta pesquisa.

Espera-se que os aspectos aqui discutidos a partir da análise do *corpus* possam contribuir para um conhecimento mais preciso do português sob uma perspectiva diacrônica, revelando um pouco mais as trilhas percorridas ao longo de sua história. Desnecessário se faz ressaltar que os dados apresentam apenas uma ponta do *iceberg*, levando-se em consideração que em diacronia é preciso saber “fazer o melhor uso dos maus dados”. De qualquer forma, evidencia-se a necessidade de construir-se uma história da língua portuguesa cada vez mais renovada.

Só voltando às origens pode-se lançar luz sobre fenômenos que ainda causam grandes embates como a variação na concordância verbal. Este trabalho busca, com suas limitações, às vezes impostas pelo próprio *corpus*, contribuir de forma significativa para uma maior compreensão desse fenômeno lingüístico, ao mesmo tempo em que pretende servir de incentivo a outros pesquisadores.

2 A VISÃO “PREÇEITIVA” À VISÃO DESCRITIVA

O embate entre descrição e prescrição há muito que tem marcado a história dos estudos lingüísticos. As reflexões sobre “como se diz” ou “como se deve dizer” acabam por gerar um novo impasse: o conceito de “norma”.

A discussão do conceito de “norma” parte da expressiva polissemia e imprecisão que esse termo recobre. Sobre essa questão, é possível enfatizar as duas faces da norma: a *norma objetiva* (conjunto de padrões habituais, costumeiros de uma comunidade de fala, ou seja, padrões observáveis na atividade lingüística de um grupo determinado: “como se diz”) e a *norma subjetiva* (conjunto de formas ideais que são impostas na comunidade de fala; sistema de valores que norteia o julgamento subjetivo do desempenho lingüístico dos falantes dentro de uma comunidade: “como se deve dizer”). O primeiro conceito relaciona-se com o que é *normal*, costumeiro, tradicional dentro de uma comunidade; já o segundo, diz respeito ao que é *normativo*, ou seja, “remete a um sistema ideal de valores que, não raro, é imposto dentro de uma comunidade” (LUCCHESI, 1994, p. 18). Dentro dessa discussão sobre o conceito de *norma*, vale ressaltar, partindo da realidade lingüística brasileira, a distinção enfatizada por Lucchesi e Lobo (1988) entre *norma padrão* (formas contidas e prescritas pelas gramáticas normativas) e *norma culta* (formas efetivamente apreendidas da fala dos segmentos plenamente escolarizados, ou seja, dos falantes com curso superior completo).

Para Lucchesi (1994), a distinção entre uma norma objetiva, observável, e uma norma subjetiva, imposta, não é tão nítida, embora possa ser útil aos objetivos teóricos e metodológicos que norteiem a investigação. O que realmente se verifica é

uma complexa interação do sistema de valores adotados por um grupo e os padrões lingüísticos observados no seu comportamento, ou seja, existe uma inter-relação entre o que é habitual e o que é imposto de forma clara e subliminar, como bem enfatiza Lucchesi (1994).

Assim, o conjunto de valores subjetivos profundamente determinados por fatores sociais, culturais e ideológicos, que é a *norma*, está intimamente relacionado às tendências e aos padrões de comportamento lingüístico que se observam numa comunidade, o que também é *norma*. Segundo Lucchesi (1994), o estudo dessa íntima relação é crucial para o entendimento da variação e das mudanças de uma dada comunidade de fala. No entanto não é essa concepção que se encontra veiculada pela tradição gramatical luso-brasileira, que sempre primou pelo caráter prescritivo.

Dito isto, faz-se necessário destacar de forma mais precisa como os gramáticos de tradição luso-brasileira têm focalizado as questões referentes à língua, sobretudo no que diz respeito ao nosso objeto de investigação, a variação na concordância verbal.

2.1 A TRADIÇÃO GRAMATICAL LUSO-BRASILEIRA

A primeira formulação prescritiva da língua portuguesa – a *Grammatica da lingua portuguesa* de João de Barros – data da primeira metade do século XVI, mais especificamente do ano de 1540, embora os primeiros documentos em português,

que estabelecem o surgimento do período arcaico, sejam de inícios do século XIII, ou da segunda metade do século XII (cf. MARTINS, 2001a). Sobre esses documentos, e conseqüente estabelecimento do período arcaico da língua portuguesa, é preciso ressaltar que

Na tradição filológica da primeira metade do século XX, estabeleceu-se que o início do período arcaico se define com o surgimento do primeiro documento oficial-régio em 1214 — *O testamento de Afonso II*, documento temporão, já que só a partir de 1255 começam a se multiplicar documentos em português na Chancelaria Real de Afonso III. Quanto a documentos particulares, seriam os mais antigos o *Auto de Partilhas* e o *Testamento de Elvira Sanches*, ambos dos fins do século XII (...), e também a *Notícia de Torto*, hoje situada entre 1212-1214. Nos meados da década de sessenta, Luís Filipe Lindley Cintra e José Avelino de Jesus, filólogo e historiador, respectivamente, reabrem a questão. Demonstram que tanto o *Auto de Partilhas* como o *Testamento de Elvira Sanches* são falsificações dos finais do século XIII. Contudo, nos finais dos anos noventa, a questão volta a ser reaberta, a partir das pesquisas de Ana Maria Martins nos fundos documentais de mosteiros, arquivados hoje no AN-TT de Lisboa. As novas pesquisas revelam documentos particulares de *scripta* conservadora. Mantém-se ainda como documento oficial mais antigo em português, de *scripta* inovadora, o Testamento de Afonso II, do qual se encontrou no arquivo da diocese de Toledo, em meados do século XX, um novo testemunho. (MATTOS E SILVA, 2002, p. 3)

Embora o problema em discussão no presente trabalho não esteja diretamente relacionado à questão histórica da periodização das línguas e, por extensão, à definição do limite inicial do período arcaico da língua portuguesa, tais embates mostram-se interessantes, uma vez que nos levam a refletir sobre o fato de que, durante o período arcaico, a língua portuguesa trilhava os caminhos de sua história sem ainda um “policiamento gramatical” explícito. Mesmo assim, considerando as pesquisas realizadas, reconhece-se que cada *scriptorium* possuía suas normas próprias de escrita, fruto de “uma tradição gráfica consolidada, uma verdadeira *scripta*”, como destaca Maia (1986, p. 14). Sobre essa questão, Marquilhas (2000) ainda argumenta que, enquanto até fins da Idade Média, o domínio da escrita se restringia a um reduzido grupo de indivíduos sujeitos a um

treino gráfico sistemático, o treino scriptológico, em períodos pós-medievais, se tem progressivamente acesso a textos lavrados por mãos pouco especializadas, às quais denomina de “mãos inábeis”.

Nesse ínterim, é justamente a *Grammatica da lingua portuguesa* de João de Barros, impressa por Luis Rodrigues em Lisboa, no ano de 1540, e dedicada ao “príncipe, nósso senhor” (BARROS, [1540] 1971, p. 292), que inaugurarás uma visão prescritiva sobre a língua, já que a *Grammatica da lingoagem portuguesa* de Fernão de Oliveira ([1536] 2000, p. 79-155), impressa por German Gallarde em Lisboa, desenvolve uma perspectiva predominantemente descritiva, ou seja, apresentam-se questões de *ortografia* (e fonética), *prosódia*, *etimologia* (morfologia) e *sintaxe*, sem um caráter normativizador.

A visão decididamente “preçeitiva”, ou prescritiva, de João de Barros pode ser observada logo no prólogo de sua *Grammatica* ao definir “gramatica e suas pártes” e ressaltar que as abordará “nam segundo convém à ordem da Gramática especulativa, mas como requére a preçeitiva” (BARROS, [1540] 1971, p. 294). Seguindo essa linha de pensamento, João de Barros já apresenta a concordância verbal como uma regra categórica, não dando margens a um uso variável, ao afirmar que “Tem máis o nome ũa concordância quando está em o caso nominativo, que [h]á-de convir com o vérbo em número e pessoa, como quando digo: eu amo” (BARROS, [1540] 1971, p. 351). Além dessa regra, João de Barros ainda resalta um caso associado à concordância com o pronome relativo na posição de sujeito. Segundo ele, “Quando o nome é relativo, [h]á-de convir com o seu antecedente em gênero, número e pessoa, como: eu amo os moços, os quáes fólgam de aprender. Este nome, moços, é do género masculino e do número plurár e da terceira pessoa” (BARROS, [1540] 1971, p. 351).

Como se pode observar nas palavras do próprio João de Barros, a relação entre o sujeito e o verbo se dá de forma categórica, ou seja, o verbo deve concordar em número e pessoa com o sujeito a que se refere. Sendo assim, sujeito na 3ª pessoa do singular deve estar associado a um verbo também na 3ª pessoa do singular. Da mesma forma que, estando o sujeito na 3ª pessoa do plural, deve a forma verbal estar na 3ª pessoa também do plural. A partir dessa posição assumida por João de Barros, verifica-se que, desde o início de sua tradição gramatical, a língua portuguesa se vinculou a um ideal prescritivo que se estenderá ao longo dos séculos e ainda se reproduz em nossos compêndios gramaticais.

M. Said Ali ([1921] 2001, p. 205), em sua *Gramática histórica da língua portuguesa*, afirma que “consiste a concordância em dar a certas palavras flexionáveis as formas de gênero, número ou pessoa correspondentes à palavra a que no discurso se referem”. Destaca-se que, através da concordância, se busca “harmonizar” termos que estabelecem entre si relações sintáticas mais próximas, por exemplo, verbo-sujeito. Em suas palavras, “o verbo, variável em número e pessoa, se empregará com a terminação que se harmonizar com o respectivo sujeito” ([1921] 2001, p. 206). Diante disso, observa-se que a posição assumida pelo gramático dá continuidade ao caráter prescritivo inaugurado por João de Barros, embora afirme que as regras explicitadas em sua gramática evidenciem traços muito gerais que

Não se aplicam a todos os casos, e, quando aproveitáveis, não podem ter a pretensão de valerem sempre com o rigor próprio das fórmulas matemáticas. Teremos de fazer, portanto, o estudo dos casos particulares mais importantes. Notaremos para alguns deles não uma, mas duas concordâncias que se contrariam, e que se justificam em épocas sucessivas. (SAID ALI, [1921] 2001, p. 206)

A *Nova gramática do português contemporâneo* de Celso Cunha e Luís F. Lindley Cintra (1985, p. 485) trata a concordância verbal como a “solidariedade entre o verbo e o sujeito”, caracterizada pela “variabilidade do verbo para conformar-se ao número e à pessoa do sujeito”. Segundo os autores, a concordância pode servir também como uma forma de evitar a repetição do sujeito, uma vez que o mesmo pode ser indicado pela “flexão verbal a ele ajustada” (p. 485). Ao lado da apresentação categórica da regra de concordância, isto é, “o verbo concorda em número e pessoa com o seu sujeito, venha ele claro ou subentendido” (p. 486), Cunha e Cintra apresentam “casos particulares”, sendo possível verificar, em alguns desses casos, um uso facultativo. Vale destacar que este uso é definido a partir do tipo de sujeito presente na sentença.

Apesar de comprovar-se uma tentativa de explicitação dos usos facultativos, buscando confronto com os usos efetivos, esta é muito “tímida”. Muitas vezes justifica-se o uso facultativo, ou seja, a opção pela forma no singular ou no plural, através dos objetivos de expressão do escritor. Por exemplo, ao explicar que, sendo o sujeito representado por uma expressão partitiva, o verbo pode ir para o singular ou para o plural, Cunha e Cintra afirmam que “a cada uma destas possibilidades corresponde um novo matiz da expressão. Deixamos o verbo no singular quando queremos destacar o conjunto como uma unidade. Levamos o verbo ao plural para evidenciarmos os vários elementos que compõem o todo” (1985, p. 488).

Em linhas gerais, Cunha e Cintra apresentam os seguintes casos particulares que podem ser considerados como facultativos², embora, em algumas situações, os gramáticos admitam a preferência por uma forma (singular ou plural):

² Os casos aqui apresentados como facultativos, com pequenas diferenças, também são os discutidos pelos outros gramáticos consultados. Além disso, apresentam-se apenas aqueles casos que têm uma relação como o objeto de nossa pesquisa, ou seja, o uso variável entre a 3ª pessoa do singular e a 3ª pessoa do plural.

- a) O sujeito, indicador de quantidade aproximada, é formado de um *número plural* precedido das expressões *cerca de*, *mais de*, *menos de* e similares – o verbo vai normalmente para o plural: Ainda assim, *restavam* (*restava*) cerca de cem viragos.
- b) O sujeito constituído por expressão partitiva (como *parte de*, *uma porção de*, *o grosso de*, *o resto de*, *metade de* e equivalentes) e um substantivo ou pronome plural: A maior parte deles já não *vai* (*vão*) à fábrica.
- c) Depois de (*um*) *dos que* (= um daqueles que), o verbo vai normalmente para a 3ª pessoa do plural: Ela passou-se para outro mais decidido, um dos que *moravam* (*morava*) no quartinho dos grandes.
- d) O sujeito é um plural aparente – se não vier precedido de artigo, o verbo fica no singular: Vassouras é que não o *esquecerá* tão cedo. Quando precedido de artigo, o verbo assume normalmente a forma plural: Os Estados Unidos, então, por sua vez, *tentam* uma demonstração espetacular.
- e) Sujeito composto posposto ao verbo – concordância com o mais próximo ou no plural: *Habita-me* (*Habitam*) o espaço e a desolação.
- f) Sujeito composto formado por infinitivos – verbo no singular: Olhar e ver *era* para mim um recurso de defesa. Caso exprimam idéias contrárias, o verbo pode ir para o plural: Em sua vida, à porfia, *Se alternam* rir e chorar.
- g) Sujeito composto ligado por *ou* e por *nem* – verbo no plural se o fato expresso pode ser atribuído aos dois núcleos: O mal ou o bem dali *teriam de vir*. Se o fato expresso só pode ser atribuído a um núcleo, o verbo fica no singular: Fui devagar, mas o pé ou o espelho *traiu-me*.

- h) A expressão *um e outro* pode levar o verbo ao plural ou, com menos freqüência, ao singular: Um e outro *tinham (tinha)* a sola rota.
- i) As expressões *um ou outro* e *nem um nem outro*, empregadas como pronome substantivo ou pronome adjetivo, exigem normalmente o verbo no singular: Nem um nem outro *havia idealizado* previamente este encontro. Só um ou outro menino *usava* sapatos. O verbo pode ir para o plural quando as expressões estiverem sendo empregadas como pronome substantivo: Nem um nem outro *desejavam* questionar.

Esses são alguns dos casos ditos facultativos, ou seja, casos em que as próprias gramáticas normativas admitem supostamente variação. No entanto há alguns questionamentos a serem feitos: está-se realmente diante de variação no sentido definido pela sociolingüística laboviana, ou seja, esse alegado uso facultativo seria um caso de regra variável? Tem o falante realmente liberdade nos usos que faz da concordância?

Parece que a “preferência” por determinada forma, como alegam os gramáticos, acaba colocando algumas restrições a essa suposta variação que se veicula nos compêndios gramaticais, razão pelo qual se pode depreender que os chamados casos facultativos não podem ser considerados como variação como definida pela sociolinguística laboviana. Ademais, a análise cuidadosa da abordagem feita por Cunha e Cintra sobre a concordância verbal revela que estes, mesmo apresentando alguns casos ditos facultativos, seguem a “visão prescritiva” inaugurada por João de Barros e tão presente na tradição gramatical portuguesa, como já fora ressaltado anteriormente.

Em sua *Moderna gramática portuguesa*, Evanildo Bechara (1999, p. 20) destaca que seu estudo gramatical “alia a preocupação de uma científica descrição

sincrônica a uma visão sadia da gramática normativa, libertada do ranço do antigo *magister dixit* e sem baralhar os objetivos das duas disciplinas”. Admite ainda estar contribuindo com “úteis sugestões ou temas de reflexão para uma proposta de melhoria da vigente nomenclatura gramatical” (BECHARA, 1999, p. 20) presente nos manuais escolares. No entanto, em relação à concordância verbal, a discussão proposta pelo gramático não se distancia dos demais consultados, mesmo ao apresentar várias observações quanto aos usos lingüísticos. Segundo ele,

A concordância pode ser estabelecida de *palavra para palavra* ou de *palavra para sentido*. A concordância de *palavra para palavra* será total ou *parcial* (também chamada *atrativa*), conforme se leve em conta a totalidade ou o mais próximo das palavras determinadas numa séria de coordenação. (...) A concordância de *palavra para sentido* se diz ainda concordância “*ad sensum*” ou silepse. (BECHARA, 1999, p. 543-4)

Bechara chama a atenção para o fato de que a liberdade de concordância “oferecida” pela língua portuguesa deve ser “cuidadosamente aproveitada para não prejudicar a clareza da mensagem e a harmonia do estilo” (1999, p. 544). Ainda destaca que

Na língua oral, em que o fluxo do pensamento corre mais rápido que a formulação e estruturação da oração, é muito comum enunciar primeiro o verbo – elemento fulcral da atividade comunicativa – para depois se seguirem os outros termos oracionais. Nestas circunstâncias, o falante costuma enunciar o verbo no singular, porque ainda não pensou no sujeito a quem atribuirá a função predicativa contida no verbo; se o sujeito, neste momento, for pensado como pluralidade, os casos de discordâncias serão aí freqüentes. O mesmo ocorre com a concordância nominal, do participio. A língua escrita, formalmente mais elaborada, tem meios de evitar estas discordâncias. (BECHARA, 1999, p. 544)

Feitas essas considerações, o gramático passa a apresentar as regras de concordância, tanto as categóricas quanto as facultativas, associando estas, muitas

vezes, a questões de estilo, prosódia, condições estéticas e restrições sintáticas. Em sua descrição do mecanismo responsável pela concordância verbal, Bechara deixa bastante evidente que, “como se dá com a relação sintática de qualquer verbo e o sujeito da oração, o normal é que sujeito e verbo (...) concordem em número” (1999, p. 558). Embora busque tratar do fenômeno da forma mais ampla possível e menos prescritiva, já que esta é a proposta inicial de sua gramática, o autor acaba assumindo de forma “implícita” uma postura normativa, o que o aproxima dos demais gramáticos, como já enfatizado.

Embora a análise, aqui apresentada, da visão adotada pela tradição gramatical sobre a concordância verbal não seja exaustiva, as questões discutidas confirmam o pressuposto de que o “olhar” é sempre o mesmo, ou seja, prescritivo, sendo este inaugurado por João de Barros, como já fora mencionado, e ratificado ao longo dos séculos. Esse comportamento assumido pela tradição gramatical luso-brasileira revela uma visão purista que desconhece a possibilidade de variação, e conseqüente mudança, a que todas as línguas estão sujeitas. Somente um novo olhar, essencialmente descritivo, possibilitará uma maior compreensão do fenômeno lingüístico em discussão, uma vez que a prescrição não revela o verdadeiro comportamento adotado pelos falantes no uso da concordância verbal.

2.2 A VARIAÇÃO NO PERÍODO ARCAICO: “OUTROS OLHARES”

A análise de textos anteriores do português, particularmente aqueles da época antes do estabelecimento das normas gramaticais durante o período clássico do século XVI, tem atestado a falta de concordância na relação entre o SN sujeito e o verbo, conforme ressaltam Naro e Scherre (1999). Essas ocorrências são normalmente encaradas como erros dos escribas ou de impressão pelos pesquisadores que, ao editarem os textos para publicação, consideram como seu dever mudar as ocorrências originais de formas verbais singulares para formas verbais plurais – marcando apenas o plural pelo ~ (til) – em contextos plurais, como prescrevem as normas gramaticais, ao invés de considerarem esses supostos erros como evidência de um uso variável. Por exemplo, no trecho “E se o padre ou a madre quiserẽ dar arra[s] por seu filho, nõ possa mays dar do dizimo (...)” (FLOS)³, a forma singular *possa*, que se encontra em um contexto plural, pode ser encarada como erro do escriba e levar o editor a mudar a ocorrência da forma para *possã*, o que revela sua falta de percepção da variação *possa~possã* para referir-se à terceira pessoa do plural.

O primeiro a documentar a variação na concordância durante o período arcaico foi Joseph Huber ([1933] 1986) em sua *Gramática do português antigo*. Ao explicitar as relações de concordância, inicialmente, Huber destaca que “é evidente que o sujeito e predicado concordam em gênero e número: *A sua face era amarella* (Euf. 363). *O padre e a madre aviã com ella grande plazer* (Euf. 358)” (p. 280).

³ Quando os exemplos apresentados foram retirados do *corpus* da pesquisa, as letras indicam o texto em que se encontram (ver lista de abreviaturas e siglas).

Observa-se então que Huber apresenta um conceito de concordância verbal que se fundamenta no olhar prescritivo que já fora discutido na seção 2.1. No entanto, conforme explicita, somente o contrário é que chama a atenção, ou seja, somente aqueles casos em que a relação entre sujeito e verbo, “predicado” nas palavras de Huber, não é obedecida é que devem merecer o olhar do investigador. Diante disso, passa-se a apresentar essas situações que “chamam a atenção”.

Embora ressalte a regra categórica, Huber apresenta exemplos em que não há a aplicação da regra prevista, buscando uma explicação para esse fenômeno. Para tanto, lança mão de critérios semânticos e sintáticos, como a sinonímia, a idéia de coletividade, a posposição do sujeito, entre outros contextos. Segundo Huber ([1933] 1986, p. 281),

Por vezes, apesar de existirem dois sujeitos, o predicado encontra-se no singular: *E per esta guisa morreo o lobo e a rraposa (Fab. 32)*. Isto verifica-se especialmente quando os sujeitos são sinónimos: *A emjuria e vergonça nom he d'aquell que a rrecebe, mays he d'aquelle que a faz (Fab. 18)*.

Apresentam-se também situações em que o sujeito está no singular, com significação plural, expressando coletividade, e o predicado no plural. Os exemplos abaixo parecem comprovar esse comportamento descrito por Huber:

- i) *E logo... se ajuntarom* deante Santo Antonio *multidom de pexes grandes e pequenos* (Milagres de Sto. Antonio 1)
- ii) — *toda a cristãydade que estã* é grã coyta (Nunes 49)

No segundo exemplo, vale ressaltar que o sujeito da oração está representado pelo pronome relativo *que*, retomando “toda a cristãydade”. Huber ainda apresenta um exemplo que considera muito interessante, visto que o mesmo sujeito, representado pelo termo *gente*, se associa a um verbo que se encontra uma vez no singular e outra no plural:

iii) *Toda gente te lamça de ssey com nojo que de ty han* (Fab. 23)

Para encerrar sua pequena descrição da variação na concordância verbal durante o período arcaico da língua portuguesa, que ocupa cerca de duas páginas de sua *Gramática do português antigo*, Huber ([1933] 1986) enfatiza que é mais raro encontrar casos em que a forma verbal esteja no singular e o sujeito no plural. Ou seja, em suas observações sobre a variação no português arcaico, Huber considera que exemplos como (iv), abaixo reproduzido, não são tão comuns, o que atesta uma tendência à regra categórica de concordância:

iv) *E hi morreo grandes gentes* (Nunes 52)

Conforme salienta, esses casos, que Huber considera raros, associam-se à posposição do sujeito. Em suas palavras: “isto parece verificar-se quando o predicado antecede o sujeito” (HUBER, [1933] 1986, p. 281). Sendo assim, justifica-se a falta de concordância através do critério sintático da posição do sujeito em relação ao verbo, no entanto há outro aspecto a ser considerado, isto é, a inacusatividade ou ergatividade⁴ do verbo em questão.

⁴ O caráter inacusativo ou ergativo de alguns verbos será mais bem discutido no capítulo de análise dos dados.

Os verbos inacusativos ou ergativos selecionam um argumento interno (sujeito) que é gerado na posição de complemento do verbo (objeto), mesmo não podendo atribuir-lhe caso acusativo, como é o caso do verbo *morreo*. Assim, o falante pode muito bem não estabelecer a concordância entre o verbo e o sujeito, uma vez que não percebe que este desempenha esta função na sentença⁵.

Embora de forma ainda “tímida”, observa-se, por parte de Huber, uma primeira tentativa de explicitar o mecanismo da concordância verbal sem a visão “preçetiva” da tradição gramatical, abrindo portas para uma reflexão metalingüística mais detalhada, coerente, na medida do possível, com uma visão que considera a variação lingüística.

Verifica-se, nos trabalhos desenvolvidos por Rosa Virgínia Mattos e Silva, outra grande contribuição aos estudos diacrônicos sobre a variação na concordância verbal no período arcaico da língua portuguesa. Ao analisar pela primeira vez a variação na concordância verbal sob uma perspectiva histórica, Mattos e Silva (1986) reflete sobre a existência de uma regra sintática variável no português arcaico. Faz-se uma análise descritiva do fenômeno lingüístico aplicada a um *corpus* do século XIV, especificamente os *Quatro Livros dos Diálogos de S. Gregório*, apresentando dados sistematicamente estruturados e numericamente avaliados para a definição de subsídios sobre o português trecentista. Mattos e Silva apresenta uma síntese dos casos de variação documentados no *corpus*, retoma a descrição de Huber ([1933] 1986), contrapondo às suas conclusões. Considerando-se os diversos contextos sintáticos examinados, observou-se que, no *corpus*:

- a) não há variação quando o sujeito coletivo está contíguo ao verbo;

⁵ Esta questão será retomada no momento em que os dados forem analisados, visto que a influência do *tipo de verbo* foi uma das variáveis lingüísticas investigadas.

- b) a presença de um qualificador plural do coletivo é um fator que favorece a concordância variável, resultando no uso da forma plural para o verbo mais próximo desse qualificador;
- c) quando o sujeito é simples, a maior distância do verbo em relação ao sujeito pode favorecer a não aplicação da regra;
- d) em relação ao sujeito composto, quando os núcleos são parassinônimos, coordenados, há um favorecimento da forma singular;
- e) sujeito simples posposto propicia a não obediência à regra enquanto que, com o sujeito simples anteposto, se costuma aplicar essa regra.

As questões postas permitem a Mattos e Silva concluir que tais informações são valiosas para um editor crítico de textos, uma vez que este poderá valer-se das mesmas para “avaliar com mais segurança quando deverá ou não interferir em grafias duvidosas que o códice apresente no que respeita à representação de P3 e P6⁶ de formas verbais” (MATTOS E SILVA, 1986, p. 97). Assim, segundo a autora, a discussão apresentada pode ser encarada como um exemplo em que “a análise descritiva detalhada da sintaxe do português arcaico pode contribuir para o estabelecimento crítico mais preciso de textos medievais portugueses” (MATTOS E SILVA, 1986, p. 97).

Em trabalho posterior, *Estruturas trecentistas: elementos para uma gramática do português arcaico*, Mattos e Silva (1989) retoma a discussão esboçada no artigo *Contribuição para a leitura crítica de textos medievais portugueses: sintaxe e grafia* (MATTOS E SILVA, 1986), chegando às seguintes conclusões:

⁶ P3 (terceira pessoa do singular) e P6 (terceira pessoa do plural).

- a) a contigüidade e maior proximidade do nome sujeito [+singular -singularidade]⁷ favorece a seleção da forma verbal de P3;
- b) quando o sujeito apresenta um sintagma nominal qualificador no plural, o verbo da relativa seguinte com o *que* referente a N (nome) + SN (sintagma nominal) qualificador da oração matriz alterna entre P3 e P6, embora estruturas desse tipo sejam raras no *corpus*⁸ analisado;
- c) na relativa subsequente ao N [+singular -singularidade] há uma preferência no *corpus* pelo verbo em P3; em enunciados de maior complexidade sintática, a encaixada mais próxima do SN de núcleo N [+singular -singularidade] apresenta-se com o verbo em P3 e, para as mais distanciadas, parece haver um favorecimento ao verbo em P6;
- d) as coordenadas com o mesmo sujeito do tipo SN de núcleo N [+singular -singularidade] em geral apresentam o verbo na mesma pessoa, mas podem ocorrer coordenadas de sujeito idêntico com verbos em pessoas distintas; nesse caso, é possível que uma maior distância na cadeia do enunciado entre o sujeito e o verbo favoreça a seleção do verbo em P6.

Essas observações levam Mattos e Silva a assumir que *a maior ou menor distância do verbo em relação ao sujeito* pode ser um fator favorecedor da concordância variável no português arcaico (verbo em P3 ou em P6). Mas a análise proposta não se esgota nos aspectos focalizados acima. A pesquisadora passa a examinar outros aspectos, cujos resultados são sumarizados abaixo:

⁷ O que Mattos e Silva chama de traço semântico [+singular -singularidade] corresponde, em termos tradicionais, ao sujeito que tem como núcleo um substantivo coletivo como em “E todo o *poboo* se *ajuntou* pêra veer” (exemplo retirado da própria autora).

⁸ O *corpus* utilizado foi a versão trecentista dos *Quatro Livros dos Diálogos de São Gregório* (documento do século XIV).

- a) com o SN simples a não contigüidade do verbo a seu sujeito pode determinar a não obediência à regra geral de concordância; da mesma forma, a concordância por sínese pode impedir a aplicação da regra;
- b) com SNs sujeitos compostos de parassinônimos, a concordância é variável, sendo mais freqüente o verbo no singular;
- c) com SNs sujeitos coordenados não-parassinônimos, é mais freqüente o verbo estar no plural; anteposição do verbo aos sujeitos permite a variação.

Observa-se, na descrição feita, uma tentativa de explicitar da forma mais exaustiva possível, levando-se em consideração as próprias “limitações” impostas pelo *corpus*, o mecanismo responsável pelo uso variável da concordância. Configura-se, assim, como uma tentativa de sistematizar “elementos para uma gramática do português arcaico”, como sugere o subtítulo da obra. Além disso, o estudo feito focalizou uma perspectiva qualitativa do fenômeno, uma vez que não se buscou fazer uma quantificação dos dados pelo menos em termos de percentagem.

Partindo de um *corpus* mais amplo que o utilizado em trabalhos anteriores, no artigo *Caminhos de mudanças sintático-semânticas no português arcaico*, Mattos e Silva (1991) reúne dados sobre pontos em variação e mudança em curso no período arcaico do português, como contribuição à história da língua portuguesa. Segundo a autora, os fatos por ela abordados são estritamente sintáticos, a saber:

- a) a variação de *ser* e *estar* e o avanço deste em estruturas atributivas;
- b) a variação de *haver* e *ter* e o avanço deste em estruturas possessivas;
- c) a variação singular/plural e a perda da marca de concordância no “particípio passado” em estruturas com *haver/ter* analisáveis como “tempo composto”;

- d) a variação singular/plural na forma e sua relação de concordância com o sujeito do verbo.

Mattos e Silva enfatiza que a dimensão cronológica das questões tratadas abrange o português literário, em prosa, trecentista e quatrocentista. A autora justifica a escolha desse tipo de texto, argumentando que

para os estudos sintático-semânticos é a documentação literária em prosa a mais rica para a aproximação histórica, estando-se, contudo, sempre consciente de que se está trabalhando sobre texto escrito, na maior parte dos casos traduzidos de outras línguas (latim, castelhano, francês...) e que as informações aí depreendidas são apenas pistas ou rastros para o que seria o português falado de então. (MATTOS E SILVA, 1991, p. 61)

Esclarecidas essas questões, passa-se à análise dos dados coletados. Na quarta parte do artigo, onde se discute a variação na concordância verbal, ressalta-se que os dados trecentistas mostram a influência, pelo menos, dos fatores *distância entre sujeito e verbo, posposição do sujeito, tipo de sujeito e semântico*, como uma interpretação singular ou plural para o sujeito, condicionando a variação. Além disso, os exemplos analisados, e exaustivamente discutidos, revelaram que a variação ocorre não apenas quando a marca de plural do verbo é o traço nasal, representado no manuscrito ou por *til* ou por *m, n*. Assim, a partir da análise de *corpus*⁹ mais amplo, como já mencionado, verifica-se que as conclusões a que chegou a pesquisadora confirmam às de trabalhos anteriores¹⁰: os dados são indícios de que já no período arcaico a posição e distância do sujeito em relação ao verbo, além de fatores semânticos, favorecem a variação na concordância verbal. No entanto deixa

⁹ O *corpus* deste trabalho é constituído dos seguintes textos: *A mais antiga versão portuguesa dos 'Quatro Livros dos Diálogos de São Gregório'*, da segunda metade do século XIV; *A lenda do rei Rodrigo*, documento da primeira metade do século XV; *A crônica de d. Pedro de Fernão Lopes*, também da primeira metade do século XV; *A linguagem da Imitação de Cristo*, texto da segunda metade do século XV; *Diálogo da Viciosa Vergonha de João de Barros*, da primeira metade do século XVI.

¹⁰ Ver os resultados apresentados em Mattos e Silva (1986, 1989), aqui já discutidos.

claro que é preciso fazer uma maior sistematização do fenômeno em questão, buscando analisá-lo em um *corpus* mais representativo, tanto no que se refere à abrangência quanto à diversidade tipológica dos textos, da tradição medieval portuguesa, como se propõe a presente pesquisa.

Diante do exposto, corroborando o objetivo perseguido por Mattos e Silva, as informações históricas podem contribuir para os estudos sincrônicos sobre variação/mudança nas regras de concordância do português contemporâneo, sendo possível, também, através da observação de processos de mudança que estão a desenvolver-se no presente, inferir como se deram no passado os processos de mudança lingüística. O que permite evidenciar a importância do *princípio da uniformidade* da *Teoria da variação e mudança*, sendo o qual “as forças que operaram no passado para produzir a mudança histórica são as mesmas que operam no presente”¹¹ (LABOV, 1975, p. 829).

As questões abordadas até aqui serão retomadas por Mattos e Silva em mais dois trabalhos: *O português arcaico: morfologia e sintaxe* (1994), em seção dedicada à *Concordância verbo-nominal*; e *A concordância verbo-nominal facultativa no português arcaico* (1998b), Comunicação apresentada no IX Congresso da ALFAL, realizado na UNICAMP.

Embora de forma bastante sintética, em virtude da proposta editorial, mais uma vez Mattos e Silva (1994) apresenta uma reflexão sobre o fenômeno da concordância no período arcaico da língua portuguesa. Partindo da primeira formulação prescritiva da gramática do português, a de João de Barros, em 1540, e retomando a variação documentada por Huber ([1933] 1986), a pesquisadora infere, a partir do conjunto de dados observados, que a variação na concordância decorria,

¹¹ “the forces which operated to produce the historical record are the same as those which can be seen operating today”.

sobretudo, de fatores semânticos e fatores sintáticos. Mattos e Silva (1994, p. 90) destaca que, na documentação sobre a variação que reuniu,

presença/ausência [de marcas de concordância] ocorre tanto com verbos em que a distinção fônica entre o singular e o plural da 3ª pessoa é mínima (apenas o travamento nasal distingue singular e plural do verbo), mas também com verbos em que a distinção mórfica é “saliente” (...).

Para Mattos e Silva, esta constatação permite defender que, nos casos em que a distinção entre a forma singular e a forma plural é mínima, ou seja, a distinção se dá apenas pelo uso do ~ (til) na forma plural, não se pode atribuir a não aplicação da regra de concordância a um lapso de escrita. A autora conclui sua breve apresentação com uma informação diacrônica: o latim falado fazia uso mais freqüente da concordância “ad sensum” do que o latim culto, pelo menos com o sujeito coletivo e sujeito composto ligado pela conjunção *cum*. Segundo a pesquisadora essa informação é significativa, uma vez que, em parte, coincide com o que foi observado na documentação por ela examinada.

Mattos e Silva (1998b)¹² rediscute *A concordância verbo-nominal facultativa no português arcaico*, destacando, inicialmente, como a questão se configura no português contemporâneo. Conforme a autora, observa-se, no PB, a coexistência de um sistema de regras que prevê, de forma idealizada, que falantes cultos deverão realizar a concordância de número entre o sujeito e o verbo da oração e outro sistema de regras facultativas que conduz a perdas das marcas formais de concordância, podendo chegar aos limites da simplificação as regras de

¹² O *corpus* utilizado nesse trabalho compõe-se, além da versão trecentista dos *Quatro Livros dos Diálogos de S. Gregório*, de documentos tanto contemporâneos como imediatamente posteriores ao texto trecentista: o *Livro das Aves*, manuscrito provavelmente copiado no mesmo tempo e no mesmo *scriptorium* que os *Diálogos*; o conjunto de *Vidas de Santos*, cópias dos meados de século XV, do *scriptorium* alcobacense, provavelmente de textos que podem ser situados no século XIV ou começos do XV; a *Vida de Santo Aleixo*, situável entre fins do século XIV e começos do XV, também alcobacense.

concordância. Em contrapartida, “afirma-se sempre que o português europeu mantém o conjunto de regras em causa de forma categórica” (MATTOS E SILVA, 1998b, p. 165). Essa situação, ou seja, a divergência sintática entre essas duas variedades da língua portuguesa, tem sua razão nas histórias particulares de cada uma, o que leva Mattos e Silva (1998b, p. 167) a colocar como centro de seu trabalho “uma volta pelo passado remoto do português, tanto europeu como brasileiro, para verificar se dele se pode depreender informação histórica que possa fornecer elementos para melhor explicitar a realidade atual”, além do próprio português arcaico.

Mattos e Silva (1998b) volta seu olhar para a depreensão dessas informações históricas sobre a concordância verbal, ressaltando que as gramáticas históricas do português, uma vez que não enfocam a sintaxe, não tratam desse fenômeno lingüístico. Em seguida, faz-se referência a Huber ([1933] 1986) que, em sua *Gramática do português antigo*, apresenta informações significativas sobre o problema no período arcaico. Chama-se a atenção para o fato de que a análise descritiva por ela desenvolvida em trabalho anterior¹³, aqui já discutido, permitiu que se confirmassem as informações de Huber, mas, além disso, que a distância do sujeito em relação ao verbo e não apenas a posposição interferia na aplicação da regra geral.

Os dados apresentados por Mattos e Silva (1998b) confirmam a existência de uma regra sintática variável, definida pela presença de dois aspectos importantes como favorecedores da mesma na fase arcaica do português: (i) possibilidade de dupla interpretação semântica do sujeito; e (ii) maior distância do sujeito em relação ao verbo. Sendo assim, argumenta-se que a concordância facultativa no período

¹³ Referência ao trabalho *Estruturas trecentistas: elementos para uma gramática do português arcaico* (1989).

arcaico decorria principalmente de fatores semânticos e sintáticos, posição já defendida em trabalhos anteriores, mas a diferença fônica mínima poderia interferir, quando relacionada a estruturas sintáticas complexas em que o verbo se apresenta distanciado de seu sujeito. A pesquisadora enfatiza que as questões levantadas evidenciam a necessidade de testá-las em um *corpus* representativo, por mais abrangente, da fase arcaica da língua portuguesa, buscando fazer uma descrição exaustiva dos contextos lingüísticos que favorecem esse fenômeno, a partir da análise de variáveis que envolvem não só fatores sintáticos e semânticos mas também morfo-fônicos.

Como se pode depreender do estudo realizado por Mattos e Silva (1998b), há algumas motivações que tornam interessante a observação desse fenômeno no período arcaico:

- i) o fato de esse período preceder ao momento em que o português já começa a se tornar normatizado¹⁴, ou seja, a partir do século XVI;
- ii) conseqüentemente, apesar de não se poder apreender como se falava no período arcaico, a sua representação estará necessariamente mais próxima da fala do que nos períodos posteriores em virtude de estar mais livre do controle normatizador, não se perdendo de vista as questões postas por Maia (1986) e Marquilhas (2000) anteriormente apresentadas (cf. nota 14).

Outros trabalhos que focalizam a variação na concordância no período arcaico foram os desenvolvidos por Anthony Julius Naro e Maria Marta Pereira Scherre: *Restrições sintáticas e semânticas no controle da concordância verbal em*

¹⁴ Não estão sendo negadas as normas de escrita que cada *scriptorium* definia, sobretudo os treinos gráficos sistemáticos, ou seja, treinos scriptológicos, que perduraram até fins do período medieval. O que se destaca aqui é o surgimento da primeira gramática prescritiva do português em meados do século XVI, a partir da qual o português começa seu processo de normatização explícita.

português (1998)¹⁵ e *Variable concord in Portuguese: the situation in Brazil and Portugal* (2000).

Naro e Scherre (1998) defendem que, embora de forma mais limitada, estágios antigos do português também têm mostrado certas áreas em variação. Ao analisarem a interação entre o traço de número e o traço [humano] do sujeito no controle da concordância em dados de documentos do português do século XIII ao XVI¹⁶, os pesquisadores argumentam que “o traço [humano] é extremamente poderoso na determinação de marcas de número no verbo” (NARO; SCHERRE, 1998, p. 67). Para eles, esse traço tanto é significativo no sentido de reter o controle da concordância quando o núcleo é [+humano], quanto no sentido de deslocar o controle da concordância para o núcleo do sintagma preposicionado (SPrep) [+humano] plural. Busca-se, então, mostrar que “o fato de maior interesse a respeito da variação da concordância verbal na história do português é que, da mesma forma que no português moderno, o traço [humano] a restringe de forma bastante acentuada” (NARO; SCHERRE, 1998, p. 62-3).

Além do traço [humano], como destacam Naro e Scherre, a análise dos casos de sujeito composto revelou a existência de muitas variáveis envolvidas no condicionamento da concordância no português antigo, além de um comportamento bastante semelhante aos resultados para o português moderno (ora para a fala, ora para a escrita), mostrando-se também relevantes as seguintes variáveis:

- a) presença e posição do sujeito em relação ao verbo;

¹⁵ Vale ressaltar que, nesse artigo, os autores, além dos dados do período arcaico, analisam dados do português falado, retirados do *Corpus Censo* do Programa de Estudos sobre o Uso da Língua (PEUL), e dados do português do Brasil escrito na década de 1990 – textos escritos por pessoas predominantemente de escolarização de nível superior, os quais passaram, em maior ou menor grau, por algum processo de revisão.

¹⁶ O *corpus* do português arcaico utilizado pelos autores compõe-se dos seguintes textos: *A Demanda do Santo Graal* (século XIII); *Diálogos de São Gregório* (século XIV); *Boosco Deleytoso* (séculos XIV e XV); *Crónica del Rei dom Joam de boa memoria por Fernão Lopes*; *Vida e Feitos de Júlio César* (século XV); *Ho Preste loam das Índias* (séculos XIV e XVI); *Crónica Geral de Espanha* de 1344 (Ms. do século XV).

- b) paralelismo discursivo (marcas precedentes em verbos que ocorrem em série);
- c) presença e posição de marcas plurais nos núcleos do sujeito composto;
- d) natureza da coordenação (*e, ou, nem*);
- e) presença de nomes próprios nos núcleos do sujeito composto;
- f) parassinonímia dos elementos coordenados;
- g) saliência fônica da oposição singular/plural;
- h) eventuais marcas de plural do complemento, quando o sujeito composto é complexo.

Os pesquisadores concluem sua análise, afirmando que, “diante das generalizações apresentadas (...), abrem-se novas perspectivas para o entendimento mais amplo da variação que envolve a concordância de número na língua portuguesa” (1998, p. 69).

Naro e Scherre (2000), por sua vez, buscando uma explicação para as origens do português brasileiro, ao realizarem um estudo em oito textos do português medieval, encontraram, conforme explicitam, mais de 200 ocorrências de formas verbais de terceira pessoa do singular em contextos em que a norma exige obrigatoriamente formas plurais de terceira pessoa. Após a codificação de cada um dos casos de ausência da marca explícita de concordância de acordo com as categorias que, segundo os pesquisadores, são estatisticamente válidas hoje no Brasil, os dados foram submetidos ao programa de regra variável VARBRUL.

A partir da análise dos dados, os pesquisadores argumentam que os fatores controladores da variação no português medieval são os mesmos que controlam a variação no PB, uma vez que “as restrições variáveis que governam o uso da

concordância não mudaram com o passar do tempo, somente mudou o peso do *input*¹⁷ (NARO; SCHERRE, 2000, p. 250).

Assim, fundamentando-se na análise do fenômeno variável, Naro e Scherre retomam a discussão sobre a natureza e origem do PB e mostram que o quadro traçado pela pesquisa se relaciona de forma mais precisa com o ponto de vista que defende a origem européia, mas também se apóia no efeito catalisador da *nativização* – passagem de uma língua não-nativa a uma língua nativa em uma comunidade de fala. Diante disso, defendem que o

português moderno do Brasil é o resultado natural da deriva secular inerente na língua trazida de Portugal, indubitavelmente exagerada no Brasil pela exuberância do contato de adultos, falantes de línguas das mais diversas origens, e da nativização desta língua pelas comunidades formadas por estes falantes.¹⁸ (NARO; SCHERRE, 2000, p. 251)

Os estudos realizados sob uma perspectiva descritiva, e que aqui foram sintetizados, revelam que estágios antigos da língua portuguesa têm mostrado que a variação na aplicação da regra categórica de concordância decorria principalmente de fatores semânticos e sintáticos, embora os fatores fônicos não sejam descartados, como, por exemplo, a influência da saliência fônica.

Observa-se, nas abordagens feitas sobre a concordância verbal no período arcaico da língua portuguesa, uma tensão entre duas visões conflitantes: *prescrição* (“como a língua deve ser”) e *descrição* (“como a língua é”). Aquela, inaugurada com

¹⁷ “the variable constraints that govern the use of agreement did not change over time; only the input weight changed”.

¹⁸ “Modern Brazilian Portuguese is the natural result of the centuries-old drift inherent in the language brought from Portugal, exaggerated, to be sure, in Brazil by extensive contact between adult language speakers of particularly diverse origins and the nativization of Portuguese by communities of such speakers”.

a primeira reflexão metalingüística¹⁹ sobre a língua portuguesa, a *Grammatica da lingua portuguesa* de João de Barros (1540), perdura em toda a tradição gramatical luso-brasileira; esta, presente em trabalhos que buscam, a partir da análise de dados e conseqüentes generalizações, abrir novas perspectivas para o entendimento mais amplo da variação que envolve a concordância de número na língua portuguesa, tem mostrado um longo caminho a percorrer – objetivo primeiro do presente estudo.

¹⁹ Vale ressaltar a *Grammatica da lingoagem portuguesa* de Fernão de Oliveira, datada de 1536, como, na verdade, a primeira reflexão metalingüística sobre o português, mas sob uma perspectiva descritiva. Assim, a gramática de João de Barros seria a primeira reflexão metalingüística de cunho prescritivo.

3 CONSIDERAÇÕES TEÓRICO-METODOLÓGICAS

Na presente pesquisa, observa-se o fenômeno da concordância verbal em textos representativos do período anterior à normatização da língua portuguesa a partir do século XVI, mais especificamente a primeira fase do período denominado arcaico, mais especificamente os séculos XIII e XIV.

Neste capítulo, serão apresentados alguns aspectos teórico-metodológicos que direcionaram a análise do fenômeno lingüístico em estudo. Para tanto, parte-se do modelo adotado – a Teoria da variação e mudança –, passando pela delimitação do problema, as relações entre o “oral” e o “escrito”, enfocando questões relativas à transmissão textual, já que se trata de um trabalho empiricamente baseado em textos escritos de fase pretérita da língua, a caracterização do *corpus*, as variáveis lingüísticas e extralingüísticas, e, finalmente, o suporte quantitativo utilizado para analisar as relações entre variável dependente e variáveis independentes.

3.1 TEORIA DA VARIAÇÃO E MUDANÇA LINGÜÍSTICA

A apresentação e discussão do fenômeno lingüístico em análise neste trabalho fundamentam-se no quadro teórico desenvolvido por Weinreich, Labov e Herzog (1968), no artigo *Empirical foundations for a theory of language change*. Este referencial teórico, conhecido sob vários rótulos – teoria da variação e mudança,

sociolingüística quantitativa, sociolingüística variacionista, sociolingüística laboviana, entre outros –, assume, como coordenada básica da reflexão sobre a mudança, a *heterogeneidade* normal da língua e, ao mesmo tempo, argumenta contra a idéia tradicional entre os lingüistas de que sistematicidade e variabilidade se excluem. Assim, considera-se que a variação é inerente a todo processo lingüístico, não sendo aleatória, mas comandada por restrições lingüísticas e não-lingüísticas, o que abre caminhos para a possibilidade de sistematizar a variação existente e própria da língua falada (no caso do presente estudo, da escrita). Depara-se, então, o historiador da língua, com uma questão crucial: as relações entre o “oral” e o “escrito”, tópico que será posteriormente retomado.

O estudo da variação sistemática decorrente dos condicionadores estruturais e sociais discute a concepção de língua enquanto sistema homogêneo, definida por si mesma. Partindo do pressuposto de que toda variação é motivada, ou seja, controlada por fatores de maneira que tal heterogeneidade se delineia sistemática e previsível, Mollica (1992) ressalta que cabe à sociolingüística investigar o grau de estabilidade ou de mutabilidade da variação, diagnosticando as variáveis que contextualizam as variantes e descrevendo seu comportamento preditivo, não perdendo de vista que essas variáveis podem ser internas ou externas à língua.

A busca pela “ordem na aparente desordem”, como propõem Weinreich, Labov e Herzog (1968), viabiliza-se através de cinco etapas de pesquisa – os cinco “problemas” da mudança lingüística (cf. TARALLO, 1990, p. 59):

- i) *constraint problem* (problema das restrições/ condicionamentos):
 “Conjunto(s) de possíveis mudanças e possíveis condições para mudanças que podem (possam) ocorrer em uma estrutura de um determinado tipo”.

- ii) *embedding problem* (problema do encaixamento): “Como as mudanças observadas estão encaixadas na matriz de concomitantes lingüísticos e extralingüísticos das formas em questão? Que outras mudanças se encontram associadas a determinadas mudanças de uma maneira não acidental? Como essas mudanças se encaixam no contexto da sociedade?”
- iii) *evaluation problem* (problema da avaliação): “Como as mudanças observadas podem ser avaliadas em termos de seus efeitos na estrutura lingüística, na eficácia comunicativa (por exemplo, na relação com a carga funcional), e na ampla gama de fatores não-representacionais envolvidos na fala?”
- iv) *transition problem* (problema da transição): “Os estágios intermediários que podem ser observados, ou que devem ser postulados, entre quaisquer duas formas de uma língua definida para uma comunidade lingüística em momentos diferentes. Como se manifesta a transição na sociedade?”.
- v) *actuation problem* (problema da implementação): “Que fatores são responsáveis pela implementação de mudanças? Por que mudanças de um traço estrutural acontecem em uma determinada língua em um dado momento, mas não em outras línguas com o mesmo traço, ou na mesma língua em outros momentos?”

Como salienta Tarallo (1990), para que uma teoria da mudança lingüística de bases preditivas seja viabilizada, torna-se preciso que esses cinco pontos, chamados de “problemas”, sejam devidamente satisfeitos. Além disso, como os problemas lingüístico-históricos não podem ser resolvidos com a mesma facilidade que questões sincrônicas, devido à fragmentação das informações contidas nos

documentos – sobretudo as de natureza fonética e social –, o *princípio da uniformidade* torna-se um dos princípios mais fortes do modelo de análise proposto por Weinreich, Labov e Hergoz (1968). Nessa proposta, “o presente e o passado aparecem como entidades em situação de contínuo de acontecimentos e não como demarcações autônomas” (TARALLO, 1990, p. 64).

Diante desse quadro teórico sinteticamente delineado, a presente pesquisa é um trabalho de lingüística descritiva e quantificada aplicada a um *corpus* constituído por textos da primeira fase do período arcaico da língua portuguesa (séculos XIII e XIV), visando a apresentar dados sistematicamente analisados e numericamente avaliados que fornecerão subsídios sobre essa fase do português e a questão da variação da concordância verbal.

3.2 DELIMITAÇÃO DO PROBLEMA

Dialetos não-padrão do português brasileiro apresentam fenômenos de concordância variável cuja origem suscita controvérsias, gerando assim, basicamente, duas correntes de pensamento: uma que defende sua natureza e origem exclusivamente européia, resultado natural da *deriva* secular inerente na língua trazida de Portugal; outra que, apoiando-se na presença maciça de pessoas de origem africana no Brasil, além das línguas indígenas, defende que as línguas desses africanos teriam influenciado o PB através de um estágio hipotético de um

pidgin ou crioulo de base portuguesa – hipótese da *crioulização prévia*²⁰. Em contrapartida, os pesquisadores portugueses são quase unânimes em declarar que as regras de concordância são categóricas em todos os dialetos de Portugal, como destacam Naro e Scherre (2000), tornando ainda mais interessante a controvérsia citada anteriormente. Considerando essa situação, a variação na concordância verbal configura-se como um dos pontos de divergência entre o PB e o português europeu (PE).

Sem dúvida, os trabalhos desenvolvidos sob a perspectiva variacionista têm revelado, no PB atual, a coexistência de um sistema de regras de concordância que prevê, de forma idealizada, que falantes cultos deverão realizar a concordância de número entre o sujeito e o verbo da oração e outro sistema de regras facultativas que conduz a perdas das marcas explícitas de concordância, podendo chegar aos limites da simplificação. Embora se apoiando no mesmo referencial teórico, ou seja, seguindo as teorias e métodos de orientação laboviana, esses estudos chegaram a conclusões diferentes quanto à aplicação dessas regras ou ausência delas no PB que, em síntese, podem ser assim delineadas:

- a) Naro (1981), trabalhando com a concordância verbal, com alfabetizando adultos do MOBREAL, propõe que no português brasileiro a variação dessas regras se configura como mudança – tanto *mudança* lenta de perda das regras, como mudança no sentido de aquisição das mesmas por aqueles portadores de dialetos simplificadores, mas expostos às regras de uso próprias às variantes do tipo “standard”;

²⁰ A designação *crioulização prévia*, embora esteja sendo aqui utilizada de forma “genérica”, abarca diferentes posturas adotadas pelos pesquisadores, envolvendo tanto uma abordagem fatorizada quanto uma mais generalizante (cf. Guy, 1986; Naro; Scherre, 1991, 1993; Tarallo, 1993; Baxter, 1995; Baxter; Lucchesi, 1997).

- b) Guy (1986), que aborda tanto a concordância nominal quanto a verbal e utiliza o mesmo *corpus* de Naro, defende que o fenômeno se configura como *variação estável* e não como mudança;

As divergências interpretativas – mudança ou variação estável – decorrem de vários aspectos do problema, discutidos e analisados, rigorosamente, segundo as teorias e métodos nos quais se apoiam os pesquisadores. Defendendo um processo de descrioulização do português popular brasileiro que leva à aprendizagem das regras de concordância a partir dos contextos “mais salientes”, Guy se coloca contra a posição defendida por Naro; este, por sua vez, argumenta que a perda das regras é natural e se inicia pelos contextos “menos salientes”, admitindo, contudo, que a mudança no sentido de aquisição que propõe pode iniciar-se pelos contextos “mais salientes”, concordando nesse sentido com Guy. Quanto à descrioulização do português brasileiro, discorda de Guy por considerar que não há evidências suficientes de que o PB se apresenta em processo de descrioulização.

A análise desses estudos mostra que há, além de um debate, um embate sobre a interpretação do fenômeno no português brasileiro contemporâneo, resultando em posturas divergentes quanto à natureza e origem do PB, como já destacado: de um lado, os defensores da idéia de que o português moderno do Brasil é o resultado de *deriva* secular inerente na língua trazida de Portugal; e do outro, aqueles que se apóiam no contato entre o português e as línguas africanas, que levou a língua trazida de Portugal a passar por um estágio hipotético de um *pidgin* ou crioulo²¹. Para acirrar o debate, “afirma-se sempre que o português

²¹ Embora tenhamos sido excessivamente concisos na apresentação do debate e/ou controvérsia sobre as origens estruturais do PB, é preciso destacar que as questões se configuram de forma mais complexa não se resumindo à embate entre os defensores da crioualização prévia e da deriva, a exemplo do conceito de transmissão linguística irregular proposto por Dante Lucchesi (UFBA).

européu mantém o conjunto de regras em causa de forma categórica” (MATTOS E SILVA, 1998b, p. 165).

Essa situação, ou seja, a divergência sintática entre essas duas variedades da língua portuguesa, tem sua razão nas histórias particulares de cada uma. Teorias lingüísticas recentes permitem afirmar que essa característica marcante da sintaxe brasileira é uma das conseqüências da direção tipológica do português do Brasil, que vem sendo classificado como “língua de proeminência tópica”, do que decorre, entre outras características de sua sintaxe, a perda da concordância. Já o português europeu, “considerado uma língua de proeminência de sujeito”, seguiria outra direção, o que reforçaria a manutenção das regras de concordância e outras características sintáticas próprias à variedade européia, como afirmam Duarte e Faria (1989).

Segundo Mattos e Silva (1998b), a questão não é tão simples assim, visto que há ainda muito a ser explorado no PE, sobretudo o estudo da sintaxe de suas variantes faladas, principalmente suas formas regionais, de falantes não ou pouco escolarizados. Além da necessidade de um estudo sistemático de fases pretéritas da língua portuguesa, objetivando verificar se esse comportamento estaria ou não prefigurado no sistema. Evidencia-se, então, a necessidade de reflexão sobre o passado remoto do português, tanto europeu como brasileiro, a fim de depreender informações de caráter histórico que possibilitem uma explicitação da situação atual do português.

O presente trabalho sobre a variação na concordância verbal na primeira fase do período arcaico da língua portuguesa inscreve-se na tentativa de fornecer subsídios para elucidar essa questão. Buscando tornar mais preciso o conhecimento do português no caminho de sua história, pretende-se “lançar luz” sobre as origens

do PB, uma vez que esse fenômeno variável tem sido usado para defender tanto a posição da deriva natural quanto à da crioulização prévia.

3.3 DO FALADO PARA O ESCRITO: A TRANSMISSÃO TEXTUAL

No artigo *Dos textos escritos à história da língua*, Clarinda de Azevedo Maia (2002) discute algumas questões da Linguística Histórica, a fim de definir linhas de atuação metodológica no que se refere à atitude do historiador da língua perante os textos escritos, tendo em vista um consistente aproveitamento dos materiais neles contidos.

“Depois de algumas décadas de domínio e de valorização do descritivismo sincrónico, o estudo da mudança lingüística volta a situar-se actualmente no centro da investigação lingüística” (MAIA, 2002, p. 231). Sem dúvida, ao longo da história da Linguística, o tratamento dado ao fenômeno da mudança sofreu numerosas reconsiderações. O estabelecimento de relações entre a mudança e variação lingüística muito contribuíram para a compreensão dos mecanismos reais da mudança lingüística e a explicação da sua origem e difusão. Evidencia-se isso ao se levar em consideração que

Ao contrário da Linguística Histórica tradicional que analisava e descrevia a mudança como um *resultado* susceptível de ser observado numa dada língua ou num conjunto de línguas da mesma família genética ao longo de amplos períodos de tempo, considera-se actualmente a mudança como um *processo* em todas as suas vertentes, desde o próprio processo de inovação lingüística, para cujo esclarecimento muito contribuiu a Psicolinguística e as Ciências de Cognição, passando pelo processo social

da difusão da inovação na comunidade até à completa mutação linguística e, simultaneamente, pelo processo de integração da inovação na estrutura da língua. (MAIA, 2002, p. 232)

Em síntese: mudança deixa de ser encarada como “resultado” e passa a ser vista como um “processo”. Sendo assim, o que importa, sobretudo, não é o “quando” mas o “como” e o “porquê” de cada um dos momentos desse processo “desde sua gestação como inovação até à consumação final” (MAIA, 2002, p. 232).

Nessa nova forma de encarar a mudança linguística, a Sociolinguística desempenhou e tem desempenhado um papel decisivo. Para Maia (2002, p. 232),

Depois de mais de meio século durante o qual os estudos linguísticos estavam centrados no sistema, a Sociolinguística recuperou a figura do falante, reintegrando-o na “cena comunicativa”, e desse modo, tornou possível caracterizar os agentes linguísticos inovadores e as motivações de fenómenos de inovação que estão na origem da mudança linguística.

Muitas foram as contribuições trazidas pelos estudos sobre a variação linguística realizados sob o domínio da Sociolinguística. As mesmas podem ser assim sintetizadas:

- conexão entre variação e mudança linguística;
- correlação sistemática entre variáveis linguísticas e variáveis sociais;
- compreensão do mecanismo real da mudança linguística e das condições sociais que impulsionaram esse fenómeno – pelo menos na fase da sua difusão social na comunidade – a partir da análise de casos documentados de mudanças em curso no interior de comunidades vivas (“tempo aparente”);

- *princípio da uniformidade* – através da observação de processos de mudança que estão a desenvolver-se no presente, é possível inferir como se deram no passado os processos de mudança lingüística;
- compreensão das características sociais dos falantes inovadores e dos fatores sociais que favorecem a velocidade de difusão da inovação ou que constituem condições de resistência à mudança;
- observação de que a realização da mudança dá origem necessariamente a fases de variação entre as formas lingüísticas mais conservadoras e as mais inovadoras.

Diante do exposto, evidenciam-se as valiosas contribuições que os estudos das mudanças em curso trouxeram para esclarecer aspectos da mudança lingüística anteriormente obscuros. Sem dúvida, tais esclarecimentos “podem, em muitos casos, ser de grande utilidade para ajudar a interpretar os materiais contidos em textos escritos de fases pretéritas da língua” (MAIA, 2002, p. 233).

Tendo em vista que a Lingüística Histórica é uma disciplina essencialmente interpretativa (“ars interpretandi”) e que cabe ao lingüista histórico selecionar, organizar e interpretar os dados ao seu dispor, ou seja, “fazer um bom uso dos maus dados”, os textos escritos em momentos pretéritos da história da língua impõem ao pesquisador problemas que decorrem de sua natureza parcial, fragmentária, escassa e de alcance sociolingüístico limitado. Assim, os materiais lingüísticos registrados nesses textos apenas podem constituir a base de um trabalho interpretativo a ser desenvolvido pelo lingüista histórico, sendo que “o estudo da história de uma dada língua equivale à interpretação da língua escrita” (MAIA, 2002, p. 235), aspecto que não pode ser ignorado. Diante disso, o complexo problema das relações entre o “oral” e o “escrito” manifesta-se como questão nuclear na

Lingüística Histórica, visto que, não dispondo de material falado de fases pretéritas, resta ao estudioso da história da língua a análise e interpretação dos materiais contidos nas fontes escritas.

Embora não seja de sua responsabilidade discutir os aspectos teóricos implicados na discussão das relações entre a língua oral e a língua escrita, o historiador da língua não pode perder de vista os principais problemas que, sob a perspectiva da teoria lingüística, se referem às relações entre o “oral” e o “escrito” e de com essa reflexão articular o problema da mudança lingüística. “Pode o investigador questionar-se sobre que mudanças da língua oral terminam na língua escrita e a partir de que momento do respectivo processo evolutivo” (MAIA, 2002, p. 236). A partir dos dados textuais, é possível reconstituir o processo evolutivo da língua escrita e, por vezes, podem tirar-se algumas conclusões sobre se também a língua falada mudou. Mas também é preciso levar em consideração que

as mudanças da língua reflectem nos textos escritos em graus distintos e de modo diverso de acordo com os géneros e tipos textuais. Daí decorre a necessidade de, ao analisar e descrever fenómenos evolutivos singulares, recolher os materiais empíricos num “corpus” representativo e diversificado sob o ponto de vista da tipologia textual. (MAIA, 2002, p. 237)

Ainda mais,

Em virtude da extrema complexidade do devir histórico da língua e de cada fenómeno de mudança linguística no qual estão implicados aspectos de diversa natureza – aspectos estruturais e aspectos sociais – *todos* os textos redigidos em tempos pretéritos ou apenas no lapso temporal abrangido pela pesquisa interessam ao investigador, quer se trate de documentos não-literários ou de textos literários. (MAIA, 2002, p. 240)

Pelo fato de serem datados e localizados, os textos não-literários oferecem grandes vantagens sobre os literários²², sendo possível, a partir deles, atender às vertentes diatópica, diacrônica e, em parte diastrática, de cada fenômeno evolutivo. Não obstante o seu caráter formular e, ainda, o fato de apresentar menor riqueza e diversidade de estruturas sintáticas, esse tipo de texto, de caráter privado, pode proporcionar informação ao nível de alguns aspectos da sintaxe.

Os textos literários, além do evidente interesse que apresentam para a história da língua literária, oferecem materiais mais ricos e variados para o estudo do léxico (incluindo aspectos de semântica histórica) e da sintaxe. No entanto, em virtude da complexidade da tradição textual da maioria dos textos literários, os materiais deles extraídos não oferecem, em geral, a mesma segurança para o estabelecimento da cronologia de cada fenômeno evolutivo da língua. Sendo assim, para que o historiador da língua possa avaliar criticamente os materiais contidos em cada texto literário (medieval) deve conhecer a sua história manuscrita.

Para épocas posteriores, a partir do século XVI, o investigador dispõe também de textos de reflexão metalingüística e meta-ortográfica. É sobretudo nas referências que gramáticos e ortógrafos fazem aos “erros” ou “vícios” que o historiador da língua pode detectar dados de importância para o conhecimento de algumas tendências evolutivas, assim como do desprestígio social que alguns fenômenos sofreram.

Sendo que as investigações de caráter histórico-lingüístico apóiam-se em materiais de base empírica procedentes da exploração de textos escritos do passado, como já salientado, torna-se imprescindível que o historiador da língua

²² Não pretendemos com isso reduzir a idéia de diversidade tipológica textual à distinção textos literários *versus* textos não-literários. Embora alguns afirmações se apresentem como categóricas em nossa explicitação, assumimos a existência de uma diversidade de textos literários quanto de textos não-literários.

conheça o modo de transmissão de cada texto que compõe o *corpus* documental selecionado para o estudo de cada fenômeno de mudança lingüística. Em se tratando de traduções, o pesquisador deve conhecer a cadeia de transmissão textual.

Esses aspectos colocam em evidência a “imperiosa necessidade de, em estudos de diacronia da língua portuguesa, ter em devida conta os aspectos relacionados com a tradição textual” (MAIA, 2002, p. 248). Assim, a constituição do *corpus* para o presente estudo buscou, na medida do possível, levar em consideração as questões aqui levantadas, como pode evidenciar-se na próxima seção.

3.4 *CORPUS*: CARACTERIZAÇÃO

Mattos e Silva (1991) ressalta que um fator significativo para a seleção de qualquer *corpus* que represente o português na sua fase arcaica é a avaliação prévia da edição com que se trabalhará, uma vez que, na tradição filológica, edições que são úteis a historiadores e estudiosos da literatura podem estar prejudicadas para estudos de línguas, já que seus editores interferem, mais ou menos, sem esclarecer, nos dados lingüísticos dos manuscritos que editam. Vale ressaltar o fato de que, ao editarem os textos para publicação, muitos pesquisadores consideram como seu dever mudar as ocorrências originais de formas verbais singulares para formas verbais plurais – marcando apenas o plural pelo ~ (til) – em contextos plurais,

como prescrevem as normas gramaticais. Diante disso, a escolha da documentação foi condicionada pelas edições acessíveis e de credibilidade, além da dimensão dos textos, sobretudo no que diz respeito à diversidade tipológica (textos literários e não-literários).

O *corpus* desta pesquisa²³ é constituído por documentos jurídicos ou notariais (oficiais, particulares e institucionais) e documentos literários em prosa de carácter religioso. Tais documentos²⁴ são aqui apresentados, fazendo-se uma breve caracterização dos mesmos.

3.4.1 Testamento de D. Afonso II

O *Testamento de Afonso II*, documento oficial-régio de inícios do século XIII (1214) tem sido considerado o documento oficial mais antigo em português, de *scripta* inovadora. Em meados do século XX, encontrou-se um novo testemunho no arquivo da diocese de Toledo. Nesta pesquisa, foram utilizadas as duas versões, ambas editadas pelo Pe. Avelino de Jesus da Costa (1979).

D. Afonso II, apesar de ter apenas 28 anos de idade e três meses de reinado, fez este testamento a fim de garantir a paz e a tranqüilidade da família e do reino, no caso de uma morte prematura, tendo em vista o seu estado debilitado de saúde. Assim, toma providências para garantir a sucessão legítima do reino.

²³ Com exceção dos documentos notariais editados por Maia (1986) e Martins (2001), os demais textos que compõem o presente *corpus* já fazem parte do Banco Informatizado de Textos do Programa para a História da Língua Portuguesa (BIT-PROHPOR).

²⁴ Evitou-se a utilização dos textos já analisados por Naro e Scherre (1999, 2000).

Embora tenham sido escritos treze exemplares, como ressalta Costa (1979), um para ficar no arquivo real e os outros para serem distribuídos por diversas entidades, apenas existe o enviado ao arcebispo de Braga (hoje arquivado no AN-TT de Lisboa) e o que se encontra em Toledo, descoberto recentemente. Tanto o testemunho de Lisboa quanto o de Toledo omitem o nome do respectivo notário. Além do mais, as variantes entre os exemplares “levam a supor que o original do testamento foi ditado para serem simultaneamente escritos vários exemplares por diferentes notários” (COSTA, 1979, p. 311).

3.4.2 Documentos notariais editados por Clarinda de Azevedo Maia

Os textos editados por Clarinda de Azevedo Maia (1986) – 168 documentos não-literários da Galícia e de Portugal – revelam o interesse da pesquisadora em investigar se haveria distinção entre a língua falada a norte e a sul do Minho. São documentos notariais (documentos de doações, compras, vendas, testamentos, inventários de pessoas falecidas, etc.) que denunciam o estado lingüístico da Galícia e do noroeste de Portugal desde o século XIII ao século XVI, de proveniência diversa, selecionados entre os fundos de vários arquivos portugueses e espanhóis.

Tendo em vista que o interesse do presente estudo não se ocupa dessa questão, ou seja, a investigação das divergências entre as variedades correspondentes às duas grandes regiões situadas a norte e a sul do Minho, aqui foram analisados 22 textos que se encontram distribuídos por duas províncias

portuguesas (do Douro Litoral e do Minho) que compõem a região de Entre-Douro-e-Minho, levando-se em consideração o período tomado como arco temporal da investigação, a primeira fase do período arcaico. Esses documentos são provenientes dos fundos documentais do Arquivo Nacional da Torre do Tombo (Lisboa) e do Arquivo Distrital de Braga.

3.4.3 Documentos notariais editados por Ana Maria Martins

Os textos editados por Ana Maria Martins (2001b), além de pertencerem ao Noroeste de Portugal, também representam a região de Lisboa. São documentos provenientes de fundos documentais de mosteiros, principalmente do mosteiro de Vilarinho (Noroeste) e do mosteiro de Chelas (Lisboa). Segundo a pesquisadora, a composição do *corpus* editado, formado por 218 documentos, distribui-se da seguinte forma:

- Mosteiro de Santo André de Ansedo (Baião) – 1 documento;
- Mosteiro de S. João Baptista de Pendorada (Marco de Canaveses) – 5 documentos;
- Mosteiro de S. Pedro de Cete (Paredes) – 4 documentos;
- Mosteiro de S. Pedro de Pedroso (Vila Nova de Gaia) – 11 documentos;
- Mosteiro de S. Salvador de Moreira (Maia) – 10 documentos;
- Mosteiro de S. Miguel de Vilarinho (Santo Tirso) – 77 documentos;
- Mosteiro de Chelas (Lisboa) – 110 documentos.

A produção dos textos editados localiza-se ao longo de três séculos, entre meados do século XIII e meados do século XVI. Além disso, alguns documentos situam-se na primeira metade do século XIII e outros, presumivelmente, em fins do século XII, constituindo a *produção primitiva portuguesa*.

Segundo Martins (2001b), os documentos do “Noroeste” representam, fundamentalmente, as regiões do Douro Litoral e Baixo Minho (até Braga); alguns documentos com origem no Noroeste da Beira Alta, dois documentos do Alto Minho, dois documentos feitos em Trás-os-Montes (embora dados pelo arcebispo de Braga). Já os documentos da região de Lisboa representam essencialmente uma zona que se estende ao longo do vale do Tejo, entre, aproximadamente, Lisboa e Santarém; uns poucos pertencem à margem sul do Tejo: Lavradio, Montijo e Setúbal. Para Martins,

Os documentos, que se distribuem de forma equilibrada no eixo temporal e por área geográfica, formam uma coleção organizada de modo a constituir uma base de trabalho para estudos de lingüística histórica, possibilitando enfoques comparativos de pendoros diversos. Juntamente com os documentos da mesma natureza publicados por Clarinda Maia de Azevedo Maia (1986) — 136 documentos da Galiza e 32 do Noroeste de Portugal — são um *contributo para o preenchimento de um espaço deficitário no domínio da Filologia Portuguesa, o da edição de textos não literários de proveniência regional diversa*. (grifo nosso) (2001b, p. 13)

Os documentos editados por Martins (2001b), testemunhos da vida econômica das instituições monásticas e dos que delas dependiam ou se moviam na sua esfera, são dos seguintes tipos: arrendamentos, aforamentos, empraçamentos, vendas, câmbios, partilhas, doações, quitações/renúncias, vedorias, posses, confirmações, procurações, recebimentos, notícias, testamentos, sentenças, demandas, relatos de contendas, avenças/composições/acordos. Tais documentos revelam a importância da *escrita* como instrumento indispensável tanto à gestão dos

bens quanto à gestão de conflitos (com dependentes, pares ou outros detentores de poder).

A edição desses textos revela documentos particulares de *scripta* conservadora que fazem recuar o limite inicial do período arcaico para fins do século XII. Na presente pesquisa, foram utilizados 70 desses documentos, sendo distribuídos nos dois séculos que são o foco deste estudo: séculos XIII e XIV.

3.4.4 Foro Real de Afonso X

O *Foro Real de Afonso X* trata-se de um dos primeiros textos não-literários escritos em português, visto que, após o *Testamento de D. Afonso II* (1214) e a *Notícia de Torto* (1214-1216), seguiu-se um longo hiato até aparecer novo documento em 1255, “data a partir da qual são já mais numerosos os textos em vernáculo, para o que muito contribuiu a acção de D. Afonso III e de D. Dinis” (FERREIRA, 1987, p. XV).

Segundo Ferreira,

Os fortes laços familiares que uniam os reinos de Portugal e de Castela no tempo do Rei Sábio (a sua filha D. Beatriz desposara o monarca português D. Afonso III) foram certamente uma das causas que mais contribuiu para um rápido conhecimento do grande labor cultural, literário e artístico que se estava operando na corte de Afonso X, bem como da reforma legislativa em que o monarca estava profundamente empenhado e procurava levar a cabo. (2001, p. 75)

Possivelmente, esses fatos impulsionaram de imediato a tradução para o português da obra legislativa afonsina, o que suscita alguns questionamentos ainda sem respostas satisfatórias: quando teria sido realizada a tradução? Por ordem de quem? Quais os motivos? Teve um caráter oficial ou privado? A inexistência de um corpo legal levaria à aplicação em Portugal, com adaptações, os textos de Afonso X? Seria necessário unificar juridicamente o reino, tendo em vista que ele já o estava em termos políticos? Embora haja uma tradição que busca responder a essas questões, a “análise interna [da obra afonsina] (...) não fornece elementos elucidativos a esse respeito” (FERREIRA, 2001, p. 76).

Da versão portuguesa do *Fuero Real* existe apenas um manuscrito no Arquivo Nacional da Torre do Tombo: trata-se do códice nº 4 do maço 6º de Forais Antigos. A edição crítica de José de Azevedo Ferreira (1987) desse manuscrito também compõe o *corpus* utilizado na investigação aqui realizada sobre a concordância verbal no português arcaico. Nas palavras do editor, o texto em questão pertence a fins do século XIII (outros historiadores da língua o localizam em inícios do século XIV), “embora não seja o original, mas provavelmente uma cópia caligráfica mais tardia do rascunho em cursiva em que foram mantidas as características lingüísticas” (FERREIRA, 1987, p. 43).

3.4.5 Orto do Esposo

Segundo Augusto Magne, em nota introdutória à edição crítica de Bertil Maler (1956), os “códices alcobacenses que contêm a jóia da prosa medieval que traz em sua portada o título um tanto romântico de *Orto do Esposo*” (MALER, 1956, p. VII) pertencem, pelo tema, forma e título, a um gênero literário de natureza religiosa e moralizadora que se desenvolveu nos países do Ocidente europeu. O número de cópias existentes nas bibliotecas medievais demonstra a imensa popularidade do *Orto* que, como os demais tratados congêneres,

esteia a cada passo a exposição teórica da doutrina com inesgotável fartura de exemplos, utilizando para isto a história profana, apólogos, a novelística oriental, moralistas filosóficos, críticos dos costumes, de mistura com o *Flos Sanctorum*, Plínio, os Padres do deserto, *bestiários* medievais, etc. (MALER, 1956, p. VII)

No que diz respeito à datação, trata-se o *Orto do Esposo* de um manuscrito de fins do século XIV. O texto utilizado nesta pesquisa é a edição crítica do manuscrito CCLXXIII da Biblioteca Nacional de Lisboa, editado por Maler (1956). Como ressalta o editor, além desse manuscrito, conhece-se outro, existente na mesma Biblioteca, com o número CCLXXIV. Em suas palavras, os “manuscritos oferecem quase exatamente o mesmo texto, reduzindo-se as diferenças praticamente a meras variantes ortográficas ou fonéticas” (MALER, 1956, p. XII).

3.4.6 Flos Sanctorum

Para a presente pesquisa, utilizou-se a edição interpretativa do *Flos Sanctorum* apresentada por Américo Venâncio Lopes Machado Filho ao Programa de Pós-Graduação em Letras e Lingüística da Universidade Federal da Bahia como trabalho final de seu doutoramento. Nas palavras de Machado Filho (2003, p. xxiii),

o *Flos Sanctorum*, [por ele editado], (...) é um documento fragmentário, de caráter hagiográfico, escrito em pergaminho, paleograficamente datável do século XIV, integrante desde 1964 do acervo da Divisão de Coleções Especiais da Biblioteca Central da Universidade de Brasília, tendo sido trazido para o Brasil, juntamente com a versão mais antiga em português dos *Diálogos de São Gregório* e o *Livro das Aves*, pelo erudito professor brasileiro Serafim da Silva Neto, no ano de 1950.

O *Flos Sanctorum* focaliza a vida e feitos de diversos santos dos primeiros séculos de difusão do Cristianismo durante a Alta Idade Média, além de exemplos de doutrinação religiosa. Como ressalta Machado Filho (2003), esse códice foi adquirido, em Portugal, pelo filólogo e professor brasileiro Serafim da Silva Neto na década de 1950 e, juntamente com os *Diálogos de São Gregório* e o *Livro das Aves*, compõem o mais antigo conjunto de textos medievais portugueses que existem no Brasil.

3.5 VARIÁVEIS LINGÜÍSTICAS E EXTRALINGÜÍSTICAS

Neste trabalho, foi observada e descrita a variação na concordância verbal na primeira fase do período arcaico da língua portuguesa (séculos XIII e XIV), a partir dos princípios da “Teoria da Variação e Mudança”, como já enfatizado, estabelecendo uma relação entre fatores sintáticos, morfo-fônicos e semânticos, ou seja, variáveis lingüísticas, além de fatores extralingüísticos como o tempo e o tipo de texto. Sendo assim, correlacionou-se o fenômeno variável em estudo, ou seja, a variável dependente *presença/ausência de marcas formais* na relação entre o SN sujeito e o verbo, às construções variáveis da estrutura lingüística, analisando a interferência dos seguintes fatores (variáveis independentes ou explanatórias):

- a) lingüísticos: realização do sujeito, posição do sujeito em relação ao verbo, caracterização semântica do sujeito, saliência fônica, tipo de verbo, indicação do plural no sujeito, língua original do texto;
- b) extralingüísticos: tempo, tipo de texto.

3.5.1 Realização do sujeito

Os estudos desenvolvidos sob modelos gerativistas têm demonstrado uma relação entre o parâmetro “pro-drop” (parâmetro do sujeito nulo) e uma significativa redução ou simplificação dos paradigmas flexionais, denunciando, por sua vez,

vínculos entre “a riqueza flexional dos paradigmas verbais de uma língua e a possibilidade de omissão do sujeito em sentenças finitas” (DUARTE, 1993, p. 107).

Diante disso, definiu-se como uma das variáveis que influencia no comportamento da variável dependente *a realização do sujeito*, uma vez que o parâmetro *pro-drop* pode ser um dos fatores que condiciona a inserção das marcas explícitas de plural, visto que, o sujeito estando “ausente”, torna-se necessário recuperá-lo através das marcas formais.

Estando essa variável subdividida em dois fatores (sujeito realizado e sujeito nulo), a expectativa é de que haja uma influência da não realização do sujeito sobre a aplicação da regra de concordância.

3.5.2 Posição do sujeito em relação ao verbo

O fator *posição do sujeito* tem se revelado como um dos fatores mais significativos nos estudos sobre a variação na concordância verbal, tanto na dimensão sincrônica como na diacrônica (cf. MATTOS E SILVA, 1986, 1989, 1991, 1994, 1998b). Assim, buscou-se verificar a influência dessa variável, distribuindo-a em quatro contextos:

- a) sujeito imediatamente anteposto ao verbo;
- b) sujeito anteposto ao verbo com um ou mais constituintes intervenientes;
- c) sujeito imediatamente posposto ao verbo;
- d) sujeito posposto ao verbo com um ou mais constituintes intervenientes.

Espera-se desse fator que a posição anteposta, considerada a posição “normal”, já que se está diante de uma língua em que é mais comum ao falante a ordem SVO, seja mais relevante, exceto num contexto em que haja elementos intervenientes entre o SN sujeito e o verbo (contexto *b*).

3.5.3 Caracterização semântica do sujeito

Scherre e Naro (1998) observaram que o traço [humano] exerce influência em dados de fala, de escrita do PB moderno e do português arcaico. Partindo dessa conclusão, nesta pesquisa, pretendeu-se verificar a influência do fator animacidade no *corpus* sob análise, controlando as variáveis [+humano] e [-humano]. A hipótese é de que o sujeito [+ humano] favorece a inserção das marcas explícitas de plural no verbo.

3.5.4 Saliência fônica

O fator *saliência fônica*, bastante conhecido na bibliografia sobre o português falado brasileiro, postula que formas mais perceptíveis, ou seja, mais salientes, são mais prováveis de serem marcadas do que as menos salientes. Em outras palavras,

as formas em que a oposição singular/plural é menos perceptível (menos saliente) possibilitam uma maior ocorrência de formas singulares em contextos plurais. Levando em consideração esse comportamento, observa-se o “estabelecimento de uma análise detalhada da hierarquia da saliência, contendo seis graus, ou níveis, de oposição fônica entre as formas singulares e plurais dos verbos” (NARO; SCHERRE, 1999, p. 7). No presente estudo, a variável saliência fônica foi definida a partir de apenas dois níveis:

- a) formas menos salientes (*auya~auyã, sabha~sabhã*), cuja diferença entre as formas revela-se apenas pelo traço nasalidade que, no *corpus* analisado, é marcado de duas maneiras: uso do ~ (til); ocorrência variável de uma consoante nasal *m~n*.
- b) formas mais salientes (*fez~fezerũ, ueo~ueeron*), em que se observa a presença de outros elementos além da nasalidade. Também foram consideradas como mais salientes as oposições que, embora se distinguem pelo traço nasalidade, tal marca recai sobre a sílaba tônica, como a oposição *sta~stã*.

Tendo em vista que diversos trabalhos sob a perspectiva sincrônica têm verificado que formas mais salientes tendem a ser mais marcadas do que as menos salientes, ou seja, oposições mais salientes, sendo mais perceptíveis, aumentam a probabilidade de ocorrência da variante marca explícita de plural, espera-se confirmar esse comportamento também no *corpus* medieval, levando em consideração o *princípio da uniformidade*.

3.5.5 Tipo de verbo

Esse grupo de fatores relaciona-se diretamente com o tipo de argumento selecionado pelos verbos para desempenharem a função de sujeito da sentença. Sendo assim, os *tipos de verbos* a serem controlados foram subdivididos em cinco categorias como explicitado abaixo:

- a) transitivos (selecionam argumento externo e interno);
- b) inacusativos ou ergativos (selecionam um argumento interno, gerado na posição de complemento do verbo, embora não seja possível atribuir-se caso acusativo);
- c) inergativos (selecionam um argumento externo, figurando na posição de especificador do verbo e, conseqüentemente, desempenhando o papel de agente da ação verbal);
- d) cópula (selecionam uma *small clause*);
- e) auxiliares/modais.

Embora seja interessante levar em consideração o comportamento dos tipos de verbos de modo geral, a atenção voltou-se, sobretudo, para o caso dos inacusativos (ou ergativos) e dos inergativos que, de acordo com a perspectiva tradicional, são classificados como intransitivos. Como os verbos inacusativos selecionam sujeitos que compartilham propriedades semânticas com os objetos diretos dos verbos transitivos diretos, espera-se uma menor ocorrência de marcas formais com esse tipo de verbo, uma vez que a intuição do falante pode analisar o sujeito da sentença como objeto direto, resultando na ausência de marcas explícitas

de plural. Em contrapartida, a expectativa em relação aos verbos inergativos é sua maior influência em relação à aplicação da regra.

3.5.6 Indicação do plural no sujeito

Este fator, associado à questão da coesão estrutural, será levado em consideração, observando-se a atuação das seguintes variantes:

- a) indicação mórfica do plural;
- b) plural indicado com numeral;
- c) noção de plural indicada pelo valor semântico do item lexical;
- d) presença de um quantificador.

Através dessa variável, espera-se avaliar as relações entre as estratégias de pluralização do sujeito e marcação formal de plural nos verbos.

3.5.7 Língua original do texto

Maia (2002) ressalta que um estudo de caráter histórico a partir de textos literários traduzidos não deve perder de vista a possível influência da língua original das obras em questão. Como, no *corpus* analisado, há um texto traduzido do latim

(*Flos Sanctorum*) e um texto traduzido do castelhano (*Foro Real de Afonso X*), uma das variáveis observadas levou em consideração a língua em que foi escrito o texto, focalizando dois fatores:

- a) textos escritos em português;
- b) textos traduzidos.

Tendo em vista a questão posta por Maia (2002) no que diz respeito às possíveis influências da língua em que foi escrito original, a hipótese é de que nos textos traduzidos haja uma ocorrência maior de marcas explícitas de plural, uma vez que, ao traduzir o texto, o tradutor tenha “trabalhado com a língua de forma mais sistemática”, principalmente no *Flos Sanctorum*, considerando que se trata de um texto traduzido de uma língua com uma norma já bem definida, o latim.

3.5.8 Tempo

A variável *tempo* foi definida em função dos textos que compõem o *corpus* da pesquisa. Tendo em vista que o foco deste estudo limita-se à primeira fase do período arcaico da língua portuguesa, ou seja, aos dois primeiros séculos que marcam a consolidação do português como língua nacional, em função do aparecimento dos primeiros textos escritos em português, observaram duas marcas temporais, levando em consideração a época em que os textos foram escritos:

- a) século XIII;
- b) século XIV.

3.5.9 Tipo de texto

Suzanne Romaine (1982), para estudos históricos em que não é possível avaliar fatores externos como idade, sexo, escolaridade, etc., propõe duas categorias principais capazes de possibilitar uma análise de variáveis externas, a saber: tipo de texto e estilo. Partindo dessa visão, nesta pesquisa, definiu-se a variável *tipo de texto*, analisando-a segundo os seguintes fatores:

- 1) Documento não-literário oficial (curto)²⁵: as duas versões do *Testamento de D. Afonso II*.
- 2) Documento não-literário oficial (longo)²⁶ – tradução do castelhano para o português: *Foro Real de Afonso X*.
- 3) Documento não-literário particular (curto)²⁷ – *documentos notariais* editados por Clarinda Azevedo Maia.
- 4) Documento não-literário particular (curto) – *documentos notariais* (de caráter institucional) editados por Ana Maria Martins.
- 5) Documento literário escrito em português: *Orto do Esposo*.
- 6) Documento literário traduzido do latim: *Flos Sanctorum*.

Embora essa variável esteja distribuída em seis fatores, considerando os textos selecionados para compor o *corpus*, a hipótese é de que os textos não-literários, em função de seu “caráter formular”, transmissão textual, dentre outros fatores, favoreçam a aplicação da regra de concordância. Em contrapartida, espera-se que os textos literários, sobretudo devido à história de sua transmissão textual, desfavoreçam a inserção das marcas explícitas de plural.

²⁵ Classificação proposta por Cintra (1963).

²⁶ Idem nota anterior.

²⁷ Documentos de Mosteiros.

3.6 SUPORTE QUANTITATIVO

Para a quantificação dos dados, utilizou-se um programa computacional de uso específico para o estudo da variação lingüística – o VARBRUL –, que tem sido o suporte estatístico dos trabalhos desenvolvidos nessa linha de pesquisa. Sendo assim, a informação quanto ao número de ocorrências que acompanha cada item analisado, bem como sobre a influência dos fatores sintáticos, morfo-fônicos, semânticos e sociais na variação, fundamenta-se na exposição dos dados ao programa de regra variável VARBRUL. Após o levantamento dos dados, os mesmos serão codificados para que o programa computacional faça o tratamento quantitativo (estatístico e probabilístico) das ocorrências das variantes lingüísticas encontradas. O último passo será a interpretação dos resultados obtidos pelo VARBRUL.

Dos programas que compõem o pacote VARBRUL, nesta pesquisa, utilizaram-se os seguintes:

- CHECKTOK – corrige os dados de entrada e gera os dados corrigidos;
- READTOK – efetua leves transformações nos dados corrigidos pelo CHECKTOK e gera novos dados com ligeiras modificações, agrupando, opcionalmente, em um só arquivo diversos arquivos corrigidos;
- MAKECELL – recebe os dados gerados pelo READTOK, preparando-os para serem executados pelo IVARB, no caso desta pesquisa;
- IVARB – programa de regra variável que faz a análise dos pesos relativos ou análises probabilísticas na forma binária;
- CROSSTAB – efetua a tabulação cruzada de duas variáveis independentes previamente estabelecidas.

O programa VARBRUL além de fazer o cálculo dos pesos relativos de cada variável independente, apresenta uma seleção estatística dos diversos grupos de variáveis analisados. Inicialmente esta seleção ocorre em função do *nível de significância* que, nesse caso, foi arbitrado em 0,05 – isso significa que há 5% de chance de que a variável tenha sido escolhida por mera flutuação estatística e não por refletir uma diferença que estatisticamente seja relevante. O *log likelihood* (cálculo de verossimilhança máxima), responsável por medir o grau de adequação entre as *probabilidades* ou pesos relativos projetados e as freqüências observadas, é o segundo elemento a influenciar a escolha de uma dada variável.

Vale destacar que um aspecto importante do programa VARBRUL é que trabalha com diversos níveis de análises, efetuando-se comparações entre os valores probabilísticos atribuídos aos fatores das variáveis. Nas palavras de Scherre,

Os diversos níveis de análises são interessantes, porque permitem verificar com precisão a interferência entre variáveis, ocasionada por codificação superposta ou enviesamento natural de variáveis. Quando não há sobreposições ou enviesamentos, não se constatam interferências e as probabilidades do primeiro nível permanecem as mesmas, ou bastante semelhantes, até o último nível de análise. (1996, p. 48-9)

Embora seja de grande importância o programa computacional utilizado na presente pesquisa, é preciso salientar que os resultados numéricos só têm valor estatístico, cabendo ao pesquisador atribuir e interpretar seu valor lingüístico. Sem dúvida, a estatística revela-se como um instrumento importante, capaz de auxiliar o pesquisador a entender um pouco mais o comportamento de fenômenos lingüísticos, no caso em questão, a variação da concordância verbal na primeira fase do português arcaico.

4 ANÁLISE DO CORPUS

Nos textos que compõem o *corpus* da presente pesquisa, encontraram-se 5.901 ocorrências de P6 que, após codificação, foram submetidas à leitura do VARBRUL. Assim, os resultados obtidos através da análise estatística dos dados variáveis serão aqui descritos e analisados a partir de tabelas com pesos relativos, porcentagens e total de ocorrências.

Ao fazer o levantamento dos dados, algumas formas verbais, embora apresentassem marcas explícitas de plural, foram desconsideradas. As situações descartadas foram as seguintes:

- a) sujeito nulo sem um sujeito explícito em outra construção no discurso precedente, como pode ser observado dos exemplos (1) a (5),

(1) – Gram maravilha acaeceu, Senhor. Huũ homem a que **dizem** Niceforo diz de si que é cristão e roga-nos muy de coração que o matemos {F55vC2} por amor daquel que Saprício negou. (FLOS)²⁸

(2) Todo omẽ que **demandarẽ** en iuyzo e depoyz que ouuyr a demanda que lhy demanda seu *contendor*, deue a responder aquello que lhy **demandam**, si ou nõ, se nõ parar ante sy algũa deffensyõ que seya cũ *dereyto per que* nõ lhy deue a responder. (FRAX)

²⁸ As letras indicam o texto do qual foram retirados os exemplos (ver lista de abreviaturas e siglas).

(3) [...] saluo hua leyra que **chamã** da Vía Noua que iaz no eydo de Curuos e hũu tereo | de leyra que **chamã** o Çidral que iaz a par das casas de Curuos por preço que o dito Domjgos lohanis cõfessou que recebera | do dito Durã Steuez [...] (DNCM)

(4) E em aquella cidade avia hũũ homẽ, que **chamauõ** Euclides, que deseiaua muyto ouuir a sabedoria de Socrates filosofo, que moraua em Athenas. (ORTO)

(5) Ca a especia que **chamã** cynamomõ he aruor de collar de ciinza e poreo significa a memoria da morte que faz o homẽ tornar ã ciinza [...] (ORTO)

b) formas verbais em que a diferença entre singular e plural é meramente gráfica, como o exemplo (6),

(6) [...] e os coraçõões som regados pera dar fruyto de todo bem e **vem** della muytos proueitos a todos. (ORTO)

c) formas verbais na voz passiva pronominal, como em (7), (8) e (9),

(7) E isto seya, saluos os dereytos e as rendas del rey que en todo tempo se **possam** demandar. (FRAX)

(8) [...] por que se **fezerõ** aquellas treuas [...] (ORTO)

(9) Eu cada d[i]a esguardo e oolho a cõgregaçõ e o ajũtamẽto das uirtudes
jntellectuaaes e dos angios, e oolho o Senhor da gloria esplandecente
sobre todos e sabeo con a mẽte ãnos ceos, parado mentes e mirando cõ
marauilha as fremusuras dos angeos que se nã **podem** contar [...] (ORTO)

Do total de ocorrências de terceira pessoa do plural que foram consideradas, 5.583 (95%) exibem marcas explícitas de plural, enquanto que 318 (5%) apresentam marca zero. Essa frequência dos dados parece favorecer a aplicação da regra de concordância, entretanto é preciso levar em consideração as limitações impostas pelo próprio *corpus*: está-se diante de textos escritos de fases pretéritas do português – e aqui se coloca a complexa relação, anteriormente discutida, entre o “oral” e o “escrito”. Embora, no geral, os dados revelem essa tendência, o levantamento de contextos que favorecem ocasionalmente a falta de concordância possibilita verificar que a regra de concordância não era categórica nem mesmo no período arcaico, como revelam os exemplos (10) a (15):

(10) – (...) E levou-me pera huĩ logar muy deleytoso, qual homem outro nã
poderia {F71vC1} achar segundo meu entendimento. Ca a luz daquel logar e
a claridade que hi havia **mostrava** a fremosura do logar. (FLOS)

(11) E se leuar *per* força sposa alhea e *ante que* a aya de ueer cũ elha
nenhũa cousa forlhy tollecta, todo *quanto* lhy for tollecto cũ elha **aiao** o
esposo e a esposa pemeyo. (FRAX)

(12) [...] o qual herdameito a my **entregou** o dito Pedro Heanez e a dita ssa mulher; o qual herdameito eu reçeiby para o dito moesteyro. (DNCM)

(13) [...] ca de Jhesu Christo **saae** e **nacem** todas as sabedorias, nõ tam solamẽte as sciencias sanctas e diuynaaes mas ajnda as sciencias dos filosofos gentis. (ORTO)

(14) [...] ca elle da muytas uezes victoria a aquelles que **creem** e **defende** a sua sacta fé [...] (ORTO)

(15) Assy como a chuua e a neue **descende** do ceo e **molha** auõdosamẽte a terra e faz dar fructo [...] (ORTO)

Além do mais, nada impede que fenômenos quase categóricos, com frequências altas ou baixas de uso de uma das variantes, possam ser submetidos à análise quantitativa.

A discussão sobre os contextos favorecedores/desfavorecedores e a influência das variáveis consideradas será feita a partir das seleções que o pacote de programas VARBRUL fez na quantificação dos dados. Em função da não ocorrência de variação nos dados extraídos do *Testamento de Afonso II*, tanto a versão de Lisboa quanto a de Toledo, apresentando-se como categóricos, os mesmos foram eliminados das rodadas a fim de ser eliminado o KNOCKOUT que geraram. Da mesma forma, alguns fatores foram amalgamados quando a intenção

era uma compreensão mais precisa sobre o comportamento das variáveis bem como a busca de resposta a algumas indagações.

As variáveis serão discutidas observando-se a ordem em que foram selecionadas pelo VARBRUL, que é, justamente, aquela considerada estatisticamente significativa para o entendimento da variação que envolve o fenômeno lingüístico em estudo. Embora sejam apresentadas as freqüências e porcentagens, são os pesos relativos que fundamentarão as reflexões sobre o fenômeno já que constituem o valor que o programa atribui a cada fator de cada variável após serem confrontados com os demais e à média de ocorrência das variantes sob análise. Sendo assim, o peso relativo não só revela a importância estatística de um fator dentro de uma determinada variável como também mostra a importância dessa variável em relação ao conjunto de variáveis definidas pelo pesquisador a partir das hipóteses formuladas.

Das nove variáveis estudadas, o programa descartou apenas uma (a variável *tempo*), selecionando as demais na seguinte ordem de relevância:

- 1) tipo de verbo;
- 2) saliência fônica;
- 3) posição do sujeito em relação ao verbo;
- 4) indicação do plural no sujeito;
- 5) caracterização semântica do sujeito;
- 6) tipo de texto;
- 7) realização do sujeito;
- 8) língua original do texto.

Como já foi ressaltado, a análise será realizada a partir dessa ordem, mas levando-se em consideração que a estatística é apenas um instrumento de grande

importância que pode auxiliar ao pesquisador no entendimento do comportamento do fenômeno lingüístico em estudo.

4.1 TIPO DE VERBO

A primeira variável selecionada pelo VARBRUL foi o *tipo de verbo*. Esse grupo de fatores relaciona-se com o tipo de argumento que os verbos selecionam para exercerem a função de sujeito na sentença. Assim, como já explicitado na seção 3.5.5, subdividiu-se essa variável em função do argumento selecionado pelo verbo. Além disso, embora a tradição gramatical considere que os verbos mono-argumentais, tradicionalmente chamados de “intransitivos”, possuem o mesmo comportamento, os dados apresentados na Tabela 1 permitem identificar algumas diferenças.

Tabela 1 – Variável *tipo de verbo*

FATOR	FREQÜÊNCIA	PESO RELATIVO
Transitivo	2.954/3.091 96%	.45
Inacusativo ou ergativo	134/228 59%	.09
Inergativo	1.154/1.190 97%	.62
Modal/auxiliar	505/537 94%	.40
Cópula	836/855 98%	.71

Os verbos classificados sob a perspectiva tradicional de intransitivos foram divididos em dois grupos: i) verbos *inacusativos* ou *ergativos*, que selecionam um argumento interno, gerado na posição de complemento do verbo (tradicionalmente chamado de *objeto*); ii) verbos *inergativos*, que selecionam um argumento externo, que se projeta como especificador do verbo – em termos gerativistas, Spec VP.

Como esperado, os verbos inacusativos são os que menos favorecem as marcas explícitas de plural em função do argumento que selecionam. Embora a porcentagem de sua ocorrência seja de 59%, o peso relativo de .09 revela claramente a baixa influência desse fator sobre a aplicação da regra. Já os inergativos apresentam uma tendência a favorecer o contrário (peso relativo de .62). Pode-se considerar que, como o sujeito selecionado pelo verbo inacusativo “exibe um comportamento de complemento verbal”, o usuário da língua acaba por não perceber as relações formais entre o verbo e o sujeito, levando assim a um aumento da não concordância.

Também é interessante observar o comportamento da cópula que exibe um peso relativo de .71, favorecendo assim a inserção das marcas explícitas de plural. Durante a codificação dos dados, foi possível observar uma grande ocorrência da forma de terceira pessoa do plural do verbo “ser” (*son~sõ*) que estabelece com a forma singular (*he*) uma distinção fônica muito “saliente”. Esses contextos mais salientes, por serem mais perceptíveis, acabam exercendo uma grande influência na escolha feita pelo usuário da língua. Então uma questão se coloca: estaria a saliência fônica influenciando os resultados obtidos em relação à atuação da cópula? A resposta a esta questão só é possível através do cruzamento entre as variáveis *tipo de verbo* e *saliência fônica*. Como os fatores “verbo cópula” e “verbo

inacusativo” mostraram-se mais relevantes, um favorecendo e o outro desfavorecendo, respectivamente, a aplicação da regra de concordância, os resultados apresentados na Tabela 2, possivelmente, permitem uma compreensão mais ampla do comportamento dessas variáveis.

Tabela 2 – Cruzamento das variáveis *tipo de verbo* e *saliência fônica*

	Verbo cópula	Verbo inacusativo
[- saliente]	328/341 96%	59/141 42%
[+ saliente]	508/514 99%	75/87 86%
TOTAL	855	228

A análise dos dados da tabela acima permite verificar a atuação dos fatores sobre a concordância explícita. Há uma ocorrência maior de cópula nos contextos em que a distinção singular/plural é mais saliente, o que acaba por interferir na aplicação da regra. Assim, o verbo cópula apresenta traços mais salientes, resultando numa influência maior na aplicação da regra. Em contrapartida, o verbo inacusativo, quando marcado pelo traço menos saliente, acaba desfavorecendo a concordância (42%), enquanto que a favorece (86%) quando apresenta o traço mais saliente.

Observando ainda a Tabela 1, constata-se que tanto os verbos transitivos quanto os modais/auxiliares não exercem grande influência sobre a variável dependente presença de marca explícita de plural, exibindo, respectivamente, pesos relativos .45 e .40.

Em síntese, a hipótese de que os verbos inacusativos exercem uma influência negativa sobre a aplicação da regra de concordância, uma vez que o falante, intuitivamente, analisa o sujeito como objeto direto, pode ser confirmada através dos dados analisados.

4.2 SALIÊNCIA FÔNICA

Naro e Scherre (2000) destacam que, no *corpus* por eles analisado, a saliência fônica da oposição singular/plural revelou-se como um fator de grande importância na aplicação da regra de concordância no português medieval da mesma forma como tem exercido relevante influência nos trabalhos que investigam a variação no PB.

Sem dúvida, a saliência fônica é uma das variáveis mais significativas, uma vez que na presente análise foi a segunda a ser selecionada pelo programa VARBRUL. Como já esclarecido, para esta variável, foram definidos apenas dois níveis: mais saliente e menos saliente. Na Tabela 3, podem ser observados os resultados referentes a essa variável.

Tabela 3 – Variável *saliência fônica*

FATOR	FREQÜÊNCIA	PESO RELATIVO
[+ saliente]	2.768/2.845 97%	.64
[- saliente]	2.815/3.056 92%	.37

Os resultados da Tabela 3 confirmam que, no *corpus* analisado, formas mais salientes, por serem mais perceptíveis, favorecem a regra, já que o peso relativo é de .64, enquanto que as formas menos salientes desfavorecem-na, tendo em vista um peso relativo de .37.

O mesmo comportamento fora observado por Naro e Scherre (2000): para as formas mais salientes, o peso relativo era de .75; já para as menos salientes, era de .33. Isso os leva a considerarem que os fatores que controlavam a variação no português medieval são os mesmos que controlam a variação no PB atual, embora para essa variedade adote-se uma “hierarquia detalhada da saliência”.

Sendo assim, os dados aqui apresentados confirmam a posição de Naro e Scherre (2000), reconhecendo-se que as diferenças, no que diz respeito ao português arcaico, relacionam-se apenas com o peso do *input*, já que os pesos relativos, de certa forma, não se contradizem. A *saliência fônica*, variável que tem sido considerada estatisticamente significativa nas análises sobre a concordância verbal no PB, mostra-se, portanto, relevante já no período arcaico da língua portuguesa.

4.3 POSIÇÃO DO SUJEITO EM RELAÇÃO AO VERBO

A terceira variável selecionada pelo VARBRUL diz respeito a um aspecto relacionado à ordem dos constituintes nas sentenças: *posição do sujeito em relação ao verbo*. Para essa variável, definiram-se quatro fatores levando em conta, além da

posição, a presença ou não de elementos entre o sujeito e a forma verbal. Em função do aspecto estrutural da língua portuguesa, os falantes tendem a considerar os SNs pré-verbais como propícios a ocuparem a posição de sujeito, enquanto que os pós-verbais evidenciam características de objetos (complementos verbais). Os dados apresentados na Tabela 4 revelam uma gradação na influência dessa variável.

Tabela 4 – Variável *posição do sujeito em relação ao verbo*

FATOR	FREQÜÊNCIA	PESO RELATIVO
Anteposto	754/789 96%	.56
Anteposto com elemento(s) interveniente(s)	1.638/1.712 96%	.54
Posposto	515/594 87%	.39
Posposto com elemento(s) interveniente(s)	128/162 79%	.20

Pelos resultados apresentados, verifica-se a influência da coesão estrutural para a compreensão do fenômeno da concordância em português arcaico. A distância e a posposição contribuem para definir esse quadro. Este comportamento também fora observado por Naro e Scherre (2000).

Verifica-se que a existência de elementos entre o sujeito e a forma verbal tem uma influência significativa. O sujeito anteposto com elemento(s) interveniente(s) favorece, se comparado com o sujeito posposto com elemento(s) interveniente(s), a explicitação das marcas formais de plural, com peso relativo de .54 contra .20. Também vale ressaltar que, embora anteposto com elemento(s) interveniente(s),

este tipo de sujeito exerce uma influência que o posposto não apresenta, uma vez que este exibe um peso relativo de .39.

Ainda visando observar o comportamento dessa variável, seus fatores foram amalgamados em apenas dois, o que veio confirmar a influência da anteposição na aplicação da regra de concordância. A Tabela 5 apresenta os resultados.

Tabela 5 – Variável *posição do sujeito em relação ao verbo* (dados amalgamados)

FATOR	FREQÜÊNCIA	PESO RELATIVO
Anteposto	2.392/2.501 96%	.55
Posposto	643/756 85%	.34

Os pesos relativos confirmam que a posição que o sujeito ocupa em relação ao verbo pode exercer influência sobre a aplicação da regra de concordância. Sujeitos antepostos favorecem a inserção de marcas explícitas, com peso relativo de .55; sujeitos pospostos desfavorecem-na, visto que a probabilidade de concordância verbal com SNs pós-verbais é menor do que com SNs pré-verbais (peso relativo de .34).

Os dados presentes no *corpus* revelam um uso mais freqüente das formas verbais singulares em contextos de sujeitos plurais quando estes estão distantes ou pospostos, confirmando assim a importância da variável *posição relativa*.

4.4 INDICAÇÃO DO PLURAL NO SUJEITO

A quarta variável selecionada pelo programa, *indicação do plural no sujeito*, possibilitou identificar a atuação das estratégias de pluralização sobre a variável dependente *presença/ausência de marcas explícitas de plural*.

Após a primeira rodada do VARBRUL, considerando que o comportamento entre os fatores “presença de quantificador” e “plural indicado pelo numeral” era muito próximo, os mesmos foram amalgamados, já que apresentavam semelhanças quantitativas.

Diante disso, a variável em questão foi subdividida em três fatores: i) indicação mórfica do plural; ii) quantificador²⁹; e iii) valor semântico do item lexical. Na Tabela 6, são apresentados os resultados obtidos a respeito dessa variável após a amalgamação.

Tabela 6 – Variável *indicação do plural no sujeito*

FATOR	FREQÜÊNCIA	PESO RELATIVO
Indicação mórfica do plural	2.291/2.393 96%	.51
Quantificador	641/686 93%	.51
Valor semântico do item lexical	15/18 83%	.11

²⁹ Esse fator corresponde aos fatores “presença de quantificador” e “plural indicado pelo numeral” explicitados na seção 3.5.6.

Os dados revelam que os fatores “indicação mórfica do plural” e “quantificador”, com peso relativo .51 cada um, exercem a mesma influência sobre a regra, embora apresentem porcentagens diferentes (96% e 93%, respectivamente); enquanto que o fator “noção de plural indicada pelo valor semântico do item lexical”, com peso relativo de .11, assume posição contrária, visto que desfavorece a aplicação da regra.

Os exemplos (16) a (18) representam alguns contextos favorecedores da variação no *corpus*, levando em consideração o fator “valor semântico do item lexical”, marcado pelos traços [+ singular, - singularidade]:

(16) E pois que todo o poboo que aa eigreja **veera vio** isto, **derom** gram louvor e gram gloria e muytas graças ao filho de Deus [...] (FLOS)

(17) E quando o todo o poboo o **vyo** fazer seu oficio na vespera assi como era custume, **houverom** ende gram prazer. (FLOS)

(18) E com este tanto prazia a todo o poboo que toda a tresteza que **houverom** pola morte do Bispo dom Fiel, tanta foy a bondade deste, que se tornou em prazer e em alegria. E por esto **davam** muytas graças a nostro senhor que lhis proveera de tã boo pastor e de tã boo senhor. (FLOS)

O contexto apresentado nos exemplos acima também já foi descrito por Huber ([1933] 1986) em sua *Gramática do português antigo (Altportugiesisches Elementarbuch)*. Embora as ocorrências sejam poucas, são dados relevantes, pois

demonstram um uso lingüístico marcado pelo valor semântico em detrimento da prescrição formal, estrutural.

4.5 CARACTERIZAÇÃO SEMÂNTICA DO SUJEITO

O fator *animacidade* tem-se revelado, nos estudos sincrônicos, de grande importância para explicar a variação na concordância verbal. Da mesma forma, Naro e Scherre (2000) atestam a influência dessa variável para o português arcaico. Na presente pesquisa, esta variável também se mostrou estatisticamente relevante, tendo em vista que foi a quinta selecionada.

No *corpus* aqui analisado, observou-se a influência da *caracterização semântica do sujeito* a partir de dois fatores: [+ humano] e [- humano]. A tabela abaixo apresenta os resultados dessa variável.

Tabela 7 – Variável *caracterização semântica do sujeito*

FATOR	FREQÜÊNCIA	PESO RELATIVO
[+ humano]	4.747/4.942 96%	.53
[- humano]	836/959 87%	.35

Realmente, o traço [humano] do sujeito exerce influência sobre a aplicação da regra apresentando um peso relativo de .53, como pode ser observado no *corpus*.

Em função disso, essa foi a quinta variável estatisticamente relevante para o programa VARBRUL, como mencionado anteriormente.

Ainda observando a Tabela 7, verifica-se que o traço [- humano] desfavorece a inserção de marcas formais explícitas de plural, com peso relativo de .38. Naro e Scherre (2000) apresentam os seguintes pesos relativos para esses fatores: .58 para sujeitos humanos plurais e .27 para sujeitos não-humanos. Embora os valores sejam diferentes, o comportamento da variável é análogo.

Assim, os dados revelam que sujeitos humanos plurais são mais prováveis de influenciarem marca de concordância plural explícita do que sujeitos não-humanos, ou seja, o traço [+ humano] manifesta-se como extremamente relevante na determinação de marcas de número no verbo.

4.6 TIPO DE TEXTO

Trabalhos sob uma perspectiva diacrônica, fundamentados em textos escritos de fases pretéritas do português, ou outra língua qualquer, constituem-se em verdadeiros problemas para o pesquisador – principalmente quando se pretende fazer um controle de variáveis sociais que envolvem o quando, o onde, o quem, o para quem, dentre outros aspectos de ordem extralingüística. Trata-se, realmente, como bem destaca Mattos e Silva (2004, p. 109), retomando Labov, da “arte de fazer o melhor uso de maus dados”.

A variável *tipo de texto*, subdividida em função da diversidade tipológica dos textos que compõem o *corpus*, seguindo a orientação de Romaine (1982), foi proposta a fim de possibilitar uma análise de variáveis externas, embora manuseando um elemento extremamente lingüístico: o texto.

Sendo esta a sexta variável estatisticamente relevante para o VARBRUL, os resultados encontram-se na Tabela 8 abaixo.

Tabela 8 – Variável *tipo de texto*

FATOR	FREQÜÊNCIA	PESO RELATIVO
DNAM	639/654 96%	.72
DNCM	191/198 86%	.56
FRAX	1.298/1.364 95%	.50
FLOS	2.765/2.86 95%	.51
ORTO	584/683 86%	.25

Observa-se que o *Orto do Esposo*, texto literário escrito em português, desfavorece a aplicação da regra com um peso relativo de .25, enquanto outro texto também literário, mas traduzido do latim, o *Flos Sanctorum*, não exerce nenhuma influência sobre ela, tendo em vista o peso relativo de .51. Comportamento análogo pode ser observado através dos pesos relativos atribuídos aos *textos notariais (de caráter particular)* editados por Clarinda de Azevedo Maia, que apresentam peso relativo de .56; e ao texto não-literário oficial, traduzido do castelhano, o *Foro de Afonso X*, peso relativo de .50. Em contrapartida, os *documentos notariais de fundos*

documentais de Mosteiros editados por Ana Maria Martins favorecem a aplicação da regra com peso relativo de .72.

A fim de compreender melhor as relações entre textos literários e não-literários, submeteram-se novamente os dados ao VARBRUL, amalgamando os fatores dessa variável e transformando-os em apenas dois: textos literários e textos não-literários. Como na primeira rodada o programa tinha descartado a variável “língua de origem do texto”, achou-se por bem fazer uma nova exposição dos dados ao programa não levando em conta se o texto era traduzido ou escrito em português. Na Tabela 9, a seguir, estão os resultados obtidos:

Tabela 9 – Variável *tipo de texto* (dados amalgamados)

FATOR	FREQÜÊNCIA	PESO RELATIVO
Textos não-literários	2.128/2.216 96%	.60
Textos literários	3.349/3.579 94%	.44

Como era de se esperar, os textos literários não exercem influência positiva sobre a regra já que apresentam peso relativo de .44, enquanto que os textos não-literários, exibindo peso relativo (.60), favorecem as marcas explícitas de plural.

Embora as probabilidades fiquem em torno de .50, o comportamento que os textos não-literários denunciam em relação à aplicação da regra de concordância não é o mesmo apresentado pelos literários.

Essa diferença de comportamento entre textos literários e textos não-literários relaciona-se, sobretudo, com a transmissão textual. Enquanto os textos não-literários não apresentam, do ponto de vista da transmissão textual, certa complexidade, já

que são textos originais, escritos em cartórios, por tabeliães que podem ser identificados, embora não seja possível, em alguns casos, fazer uma biografia precisa de cada um deles; os textos literários apresentam problemas em relação ao número de cópias, datação, copistas, localização, enfim uma complexa transmissão textual.

Observou-se também que, nos textos notariais, carregados de fórmulas e construções cristalizadas, a maioria das ocorrências apresenta estruturas como as que podem ser observadas nos exemplos (19) a (23), o que corrobora para uma tão baixa freqüência dos contextos favorecedores da ausência de marcas:

(19) **Sabbham** quãtos este testemoỹ o **uiré** e **léér ouuiré** que donna Sancha Esteuaiz, [...] (DNCM)

(20) *Teftemoyas que presentes* **foro**: frey Pedro – 18 Steuë ez, frey Afonso da ordin dos preega-dores; Femam da Veiga e eu Johã – 19 Perez, tabaliõ de Miragaya, [...] (DNCM)

(21) Os que a esto **forõ** presentes: BééytoPeriz, alfaiame, Steuã Miguéez, Martin F(.....)dj, al – 17 fayates de Bragáá, e Johã Dominguíz, clerigos do d̄to juyz. (DNCM)

(22) **Conhoscám** quantos este prazo **uiré** e **léér ouuiré** que eu Stevam anes Abbade do monsteiro de Cety / 2 e o Priol e cõuento desse logar. ffazemos [...] (DNAM)

(23) **Sabhã** todos que en pressença de m̃ ffrancisco gíraldez publico tabaliõ de Gujmarães e das testemunhas que Adeantesom scritas ena crasta / 2 de santa Maria [...](DNAM)

Diante disso, os dados revelam que a variável *tipo de texto* desempenha um papel importante em relação ao fenômeno estudado; contribuindo, os *textos não-literários*, para a aplicação da regra de concordância, enquanto que os *literários*, fruto de uma transmissão textual mais complexa, acabam por favorecer a ausência de marcas explícitas. Vale destacar que, embora tenhamos “reduzido” a diversidade de textos em literários e não-literários, atestamos a complexidade da questão.

4.7 REALIZAÇÃO DO SUJEITO

A sétima variável estatisticamente relevante para o programa VARBRUL relaciona-se com o parâmetro *pro-drop*. Trata-se da realização ou não do sujeito e a sua influência sobre a variável dependente em estudo, cujos resultados podem ser observados na Tabela 10:

Tabela 10 – Variável *realização do sujeito*

FATOR	FREQÜÊNCIA	PESO RELATIVO
Sujeito realizado	3.031/3.253 93%	.22
Sujeito nulo	2.550/2.646 96%	.82

É interessante observar que, embora a porcentagem entre os fatores seja aproximada – 93% de ocorrências de sujeito realizado e 96% de sujeito nulo –, o comportamento dos fatores, observando os pesos relativos é totalmente oposto. Enquanto que as construções com sujeito nulo favorecem a aplicação da regra, com um peso relativo de .82, levando em consideração que as marcas formais asseguram a retomada do “sujeito ausente”, as ocorrências com sujeito realizado desfavorecem-na, peso relativo de .22.

Os dados acabam por confirmar aquilo que era intuição desde o início da pesquisa. A hipótese de que “marcas levam a marcas” e “zeros a zeros” não pode ser confirmada à luz dos dados utilizados na presente pesquisa, embora se reconheça a necessidade de uma maior investigação que envolva o uso do pronome sujeito e a inserção de marca explícita de plural. O que os dados do *corpus* aqui analisado revelam é uma necessidade de recuperar o sujeito não realizado através das marcas formais de plural; daí a relevância da variável *realização do sujeito* para o fenômeno lingüístico estudado.

4.8 LÍNGUA ORIGINAL DO TEXTO

Quando os dados foram submetidos ao programa computacional pela primeira vez, O VARBRUL tinha descartado a variável *língua original do texto*. Após a amalgamação de fatores em outros grupos, buscando uma visão mais precisa do fenômeno lingüístico sob análise e os contextos que tendem a favorecer ou desfavorecer uma ou outra variante, esta variável foi selecionada pelo VARBRUL em oitavo lugar, revelando-se, assim, estatisticamente relevante. Os resultados obtidos a partir da quantificação dos dados podem ser observados na Tabela 11 abaixo.

Tabela 11 – Variável *língua original do texto*

FATOR	FREQÜÊNCIA	PESO RELATIVO
Textos escritos em português	1.516/1.637 93%	. 42
Textos traduzidos	4.067/4.264 95%	. 53

Conforme os dados, o fator “textos escritos em português” desfavorece a aplicação da regra, com peso relativo de .42, confirmando a hipótese de que, como no processo de tradução há um uso “mais sistemático e consciente” da língua, os textos originalmente escritos em português permitem maiores “lapsos” por influência do uso efetivo. Dito de outra forma, os “textos traduzidos” apresentam uma influência maior na inserção de marcas explícitas de plural (peso relativo de .53), uma vez que, ao traduzir o texto, o tradutor tenha “trabalhado com a língua de forma mais sistemática”. É certo que os números são modestos, mas não se pode deixar de

considerar os limites que o *corpus* impõe, sobretudo por estarmos tratando de um arco temporal muito distante do presente uso da língua.

5 CONSIDERAÇÕES FINAIS

A variação na concordância verbal em português, um dos fenômenos lingüísticos mais estudados atualmente, não era aleatória nem mesmo na primeira fase do período arcaico da língua portuguesa (séculos XIII e XIV), mas definida a partir de restrições lingüísticas e, possivelmente, sociais, embora nem sempre seja possível o historiador da língua controlar variáveis desta natureza.

Da mesma forma, essa variação não pode ser considerada como lapso dos escribas nem mesmo nos contextos de menor saliência fônica em que a oposição singular/plural se dá apenas pela nasalização da vogal final. Se nesses contextos pode-se considerar a falta de concordância um “erro dos escribas”, que postura adotar quando se está diante de contextos mais salientes, como a oposição *son/sõ~he*, entre outras? Não confirmam esses supostos “erros” a existência de um uso variável? Não seriam esses “erros” momentos em que o escriba deixa escapar o uso efetivo, sem o caráter idealizador da norma, embora esta só comece a ser definida a partir do século XVI, com a *Gramática* de João de Barros? Certamente que essas questões precisam ser retomadas, a fim de tornar mais preciso o conhecimento sobre a formação histórica do português. Com este trabalho, abrem-se novas perspectivas para o entendimento mais amplo da variação que envolve a concordância verbal de número na língua portuguesa.

Esta pesquisa permitiu reafirmar que a variação na concordância verbal não é um fenômeno restrito ao português moderno. Os dados retirados do *corpus* que foi escolhido para constituir esta investigação demonstram que existe uma sistematicidade na aplicação da regra de concordância que, na maioria dos casos,

coincide com os contextos que a favorecem no PB e que têm sido investigados desde o trabalho pioneiro de Lemle e Naro (1977).

Embora o número de ocorrências de P6 sem marcas explícitas fosse muito baixo, no *corpus* analisado, os dados possibilitaram uma sistematização das variáveis que interferem no fenômeno em estudo.

A investigação permitiu verificar a relevância de oito variáveis independentes sobre a variável *presença/ausência de marcas explícitas de plural na relação entre o SN sujeito e o verbo*: tipo de verbo, saliência fônica, posição do sujeito em relação ao verbo, indicação do plural no sujeito, caracterização semântica do sujeito, tipo de texto, realização do sujeito e língua original do texto.

Constatou-se que as estruturas com verbos inacusativos não favorecem a aplicação da regra de concordância devido ao tipo de argumento selecionado por esses verbos: “um sujeito com características de objeto” (em termos gerativistas, um argumento interno). Já os verbos inergativos tendem a influenciar positivamente a regra, já que selecionam um argumento externo. Os verbos transitivos não exercem influência sobre a inserção das marcas explícitas de plural, devido à seleção de dois argumentos: um externo, que desempenha a função de sujeito e desencadeia a aplicação da regra; e um interno, que exerce a função de objeto. O verbo cópula, por sua vez, exerce uma grande influência sobre a regra, conjugado ao material fônico.

O princípio da *saliência fônica* demonstra, desde o período arcaico, uma influência sobre a aplicação da regra de concordância. Os contextos mais salientes favorecem a explicitação das marcas formais de plural, enquanto que formas menos salientes desfavorecem-na. Este é o mesmo comportamento observado no português moderno, demonstrando assim que “as forças que atuam no presente são as mesmas que atuaram no passado”.

A *posição relativa do sujeito* constitui uma variável significativa, uma vez que posição e distância que o sujeito ocupa em relação ao verbo favorece ou desfavorece a aplicação da regra de concordância. Sujeito anteposto tende a favorecer a inserção das marcas de plural no verbo.

Outro grupo de fatores que desempenha um papel significativo é a *caracterização semântica do sujeito*. O traço [+ humano], da mesma forma que no PB, tem se revelado como um fator responsável pela aplicação da regra. Sujeito [- humano] tende a favorecer a marca zero de concordância. Essa variável, assim como a variável *indicação de plural no sujeito*, revela a influência dos aspectos semânticos na aplicação da concordância verbal.

A complexidade da transmissão textual dos textos literários faz com que estes desfavoreçam a aplicação da regra de concordância, enquanto que os textos não-literários, com suas estruturas formulares e transmissão textual mais precisa, acabam por favorecer a exibição das marcas formais de plural.

A investigação também permitiu verificar que o parâmetro *pro-drop* manifesta-se como estatisticamente relevante. Enquanto que sujeitos realizados desfavorecem a aplicação da regra, sujeitos não realizados tendem a influir na explicitação de marcas de plural no verbo. No *corpus* analisado, esse comportamento está relacionado com a necessidade de recuperar o sujeito ausente através das marcas formais.

Finalmente, o fato de o texto ser ou não traduzido demonstra um papel significativo. Textos escritos em português tendem a desfavorecer a aplicação da regra de concordância. Tendo em vista que o trabalho de tradução envolve uma reflexão metalingüística, um “trabalhar sobre a língua”, os textos traduzidos exibem uma influência significativa sobre a regra.

Sem dúvida, a variação na concordância verbal no período arcaico da língua portuguesa não é aleatória, como já fora enfatizado, mas motivada por fatores de ordem semântica, sintática e morfo-fônica, além de fatores extralingüísticos, como a natureza dos textos. Esta investigação representa o passo inicial para a compreensão desse fenômeno lingüístico sob a dimensão diacrônica, visto que ainda é preciso estender essa reflexão à segunda fase do período arcaico e, quem sabe, a períodos posteriores.

Embora o peso do *input* seja menor, o comportamento das variáveis é o mesmo observado no português contemporâneo, revelando a importância do princípio da uniformidade para a compreensão do fenômeno variável. A observação e análise dos dados retirados do *corpus* possibilitaram generalizações sobre o fenômeno da concordância verbal, levando em consideração as próprias limitações que estudos diacrônicos impõem ao pesquisador. Somente uma observação sistemática possibilitará entrever vias para a compreensão da constituição histórica do português. Certamente, os dados aqui apresentados, e reflexões que suscitaram, revelam apenas uma ponta do *iceberg*.

REFERÊNCIAS

BARROS, João de ([1540] 1971). *Grammatica da lingua portuguesa*. Edição de M. L. BUESCU. Lisboa: Faculdade de Letras.

BAXTER, Alan N. (1995). Transmissão geracional irregular na história do português brasileiro – divergências nas vertentes afro-brasileiras. *Revista Internacional de Língua Portuguesa*, n. 14, p. 72-90, dez. 1995.

BAXTER, Alan N.; LUCCHESI, Dante (1997). A relevância dos processos de pidginização e crioulização na formação da língua portuguesa no Brasil. *Estudos: lingüísticos e literários*, Salvador: UFBA, n. 19, p. 65-83, mar. 1997.

BECHARA, Evanildo (1999). *Moderna gramática portuguesa*. 37. ed. rev. e ampl. Rio de Janeiro: Lucerna.

CINTRA, Luís Filipe Lindley (1963). Les anciens textes portugais non littéraires: classement et bibliographie. *Revue de Linguistique Romane*, Société de Linguistique Romane, Centre National de la Recherche Scientifique, 77, p. 40-58.

COSTA, Pe. Avelino de Jesus da (1979). Testamento de D. Afonso II, de 1214. In: COSTA, Pe. Avelino de Jesus da. *Os mais antigos documentos escritos em português: revisão de um problema histórico-lingüístico*. Coimbra: Faculdade de Letras da Universidade de Coimbra. p. 307-21.

CUNHA, Celso; CINTRA, Luís Filipe Lindley (1985). *Nova gramática do português contemporâneo*. 2. ed. Rio de Janeiro: Nova Fronteira.

DUARTE, Maria Eugênia Lamoglia (1993). Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: ROBERTS, Ian; KATO, Mary (orgs.). *Português brasileiro: uma viagem diacrônica*. Campinas: Editora da UNICAMP. p. 107-25.

DUARTE, Inês; FARIA, Isabel (1989). O paradoxo da variação: aspectos do português. *Revista Internacional de Língua Portuguesa*. n. 1, p. 21-7.

FERREIRA, José de Azevedo (1987). *Afonso X: Foro Real*. Lisboa: Instituto Nacional de Investigação Científica.

FERREIRA, José de Azevedo (2001). A obra legislativa de Afonso X em Portugal. In: FERREIRA, José de Azevedo *Estudos de história da língua portuguesa: obra dispersa*. Braga: Universidade do Minho/Centro de Estudos Humanísticos. p. 75-90.

GALVES, Charlotte (2001). *Ensaio sobre as gramáticas do português*. Campinas: Editora da UNICAMP.

GUY, Gregory R. (1986) On the nature and origins of popular Brazilian Portuguese. *Estudios sobre Espanhol da América y Lingüística Afroamericana*. Bogotá, p. 226-244.

HORA, Dermeval da (1997). Teoria da variação: uma retrospectiva. In: HORA, Dermeval da (Org.). *Diversidade lingüística no Brasil*. João Pessoa: Idéia, p.159-74.

HUBER, Joseph ([1933] 1986). *Altportugiesisches Elementarbuch*. Heidelberg: Carl Winters Universitätsbuchhandlung. (trad. portuguesa: *Gramática do português antigo*. Lisboa: Gulbenkian).

LABOV, William (1975). On the use of the present to explain the past. Estratto de *Linguistics at the crossroads*. Liviana: Editrice-Jupiter Press.

LEMLE, Mirian; NARO, Anthony Julius (1977). Competências básicas do português. *Relatório final de pesquisa apresentado às instituições patrocinadoras, Fundação MOBREAL e Fundação Ford*. Rio de Janeiro, ms.

LUCCHESI, Dante; LOBO, Tânia (1988). Gramática e Ideologia. *Sitientibus*, n. 8, Feira de Santana/UEFS, p.73-81.

LUCCHESI, Dante (1994). Variação e norma: elementos para uma caracterização sociolingüística do português do Brasil. *Revista Internacional de Língua Portuguesa*, n. 12, p. 17-28.

MACHADO FILHO, Américo Venâncio (2003). *Um 'Flos Sanctorum' do século XIV* Edições, glossário e estudo lingüístico. Tese (Doutorado em Letras) – Instituto de Letras, Universidade Federal da Bahia, Salvador.

MAIA, Clarinda de Azevedo (1986). *História do galego-português: estado lingüístico da galiza e do noroeste de Portugal desde o século XIII ao século XVI* (com referência à situação do galego moderno). Coimbra: Instituto Nacional de Investigação Científica.

MAIA, Clarinda de Azevedo (2002). Dos textos escritos à história da língua. *Conferência*. Universidade do Minho – Centro de Estudos Humanísticos. p. 231-249 (Separata).

MALER, Bertil (1956). *Orto do Esposo*. Rio de Janeiro: Ministério de Educação e Cultura/Instituto Nacional do Livro.

MARQUILHAS, M^a. Rita Braga. *A faculdade das letras: leitura e escrita em Portugal no século XVII*. Lisboa: Imprensa Nacional-Casa da Moeda, 2000.

MARTINS, Ana Maria (2001a). *Documentos portugueses do noroeste e da região de Lisboa: da produção primitiva ao século XVI*. Lisboa: Imprensa Nacional-Casa da Moeda.

MARTINS, Ana Maria (2001b). Emergência e generalização do português escrito: de D. Afonso Henriques a D. Dinis. In: MATEUS, Maria Helena Mira. (org.). *Caminhos do português*. Lisboa: Biblioteca Nacional, p. 23-71.

MATTOS E SILVA, Rosa Virgínia (1986). Contribuição para a leitura crítica de textos medievais portugueses: sintaxe e grafia. *Actes du Colloque Textuelle Portugaise*. Paris: Gulbenkian, p. 85-98.

MATTOS E SILVA, Rosa Virgínia (1989). A articulação do sintagma nominal sujeito e do sintagma verbal: concordância. In: MATTOS E SILVA, Rosa Virgínia. *Estruturas trecentistas: elementos para uma gramática do português arcaico*. Lisboa: IN-CM, p. 488-507.

MATTOS E SILVA, Rosa Virgínia (1991). Caminhos de mudanças sintático-semânticas no português arcaico. *Cadernos de Estudos Lingüísticos*. Campinas, 20, p. 59-74.

MATTOS E SILVA, Rosa Virgínia (1994). *O português arcaico: morfologia e sintaxe*. São Paulo: Contexto.

MATTOS E SILVA, Rosa Virgínia (1998a). Idéias para a história do português brasileiro: fragmentos para uma composição posterior. In: CASTILHO, Ataliba Teixeira de. (org.). *Para a história do português brasileiro*. v 1. Primeiras idéias. São Paulo: Humanitas/FAPESP.

MATTOS E SILVA, Rosa Virgínia (1998b). A concordância verbo-nominal facultativa no português arcaico. *Atas do IX Congresso da ALFAL*. v. IV. Campinas: UNICAMP, p. 165-175.

MATTOS E SILVA, Rosa Virgínia (2002). Novos indicadores para os limites do português arcaico. *Comunicação ao Congresso Nacional do GELNE*. Fortaleza, 2002.

MATTOS E SILVA, Rosa Virgínia (2004). *Ensaio para uma sócio-história do português brasileiro*. São Paulo: Parábola.

MOLLICA, Maria Cecília (1992). Sociolingüística: conceituação e delimitação. In: MOLLICA, Maria Cecília (Org.). *Introdução à sociolingüística variacionista*. Rio de Janeiro: Cadernos Didáticos UFRJ, p. 13-15.

MONGUILHOTT, Isabel de Oliveira e Silva; COELHO, Izete Lehmkuhl (2002). Um estudo da concordância verbal de terceira pessoa em Florianópolis. In: VANDRESEN, Paulino (org.). *Varição e mudança no português falado na região sul*. Pelotas: EDUCAT. p. 189-216.

NARO, Anthony Julius (1981). The social and structural dimensions of a syntactic change. *Language LSA*, v. 57, n.1, p. 63-98.

NARO, Anthony Julius; SCHERRE, Maria Marta Pereira (1991). Variação e mudança lingüística: fluxos e confluxos na comunidade de fala. *Cadernos de Estudos Lingüísticos*, n. 20, p. 9-16.

NARO, Anthony J.; SCHERRE, Maria Marta Pereira (1993). Sobre as origens do português popular do Brasil. *D.E.L.T.A.*, n. 9 (número especial), p. 437-455.

NARO, Anthony Julius; SCHERRE, Maria Marta Pereira (1997). A concordância de número no português do Brasil: um caso típico de variação inerente. In: HORA, Dermeval da (Org.). *Diversidade lingüística no Brasil*. João Pessoa: Idéia, p. 93-114.

NARO, Anthony J.; SCHERRE, Maria Marta Pereira (2000). Variable concord in Portuguese: the situation in Brazil and Portugal. In: WCWHORTER, John (org.). *Language change and language contact in pidgins and creoles*. John Benjamins, Amsterdam/Philadelphia. p. 235-255.

OLIVEIRA, Fernão de ([1536] 2000). *Gramática da linguagem portuguesa*. Edição crítica, semidiplomática e anastática de A. TORRES e C. ASSUNÇÃO. Lisboa: Academia das Ciências de Lisboa.

OMENA, Nelize Pires de (1992). Influências morfo-sintáticas e semânticas. In: MOLLICA, Maria Cecília (Org.). *Introdução à sociolingüística variacionista*. Rio de Janeiro: Cadernos Didáticos UFRJ, p. 47-56.

ROMAINE, Suzanne (1982). The extralinguistic variables: methods for the reconstruction of language in its social context. In: ROMAINE, Suzzane. *Socio-historical linguistics: its status and methodology*. Cambridge: Cambridge University Press, p.105- 138.

SAID ALI, M. ([1921] 2001). *Gramática Histórica da Língua Portuguesa*. 8. ed. rev. atual. São Paulo: Melhoramentos/ Brasília: Editora Universidade de Brasília.

SCHERRE, Maria Marta Pereira (1992). *Introdução ao pacote VARBRUL para microcomputadores*. Rio de Janeiro, UFRJ, Faculdade de Letras. Departamento de Lingüística e Filologia. Projeto de Estudo sobre os uso da língua (PEUL), 55 p.

SCHERRE, Maria Marta Pereira; NARO, Anthony Julius (1998). Restrições sintáticas e semânticas no controle da concordância verbal em português. *Fórum lingüístico*, Florianópolis, UFSC, v. 1, n. 1, p. 45-71, jul./dez.

SCHERRE, Maria Marta Pereira (1994). Aspectos da concordância de número no português do Brasil. *Revista Internacional de Língua Portuguesa*. n. 12, p. 37-49.

SCHERRE, Maria Marta Pereira (1996). Pressupostos teóricos e suporte quantitativo. In: SILVA, Giselle Machline de Oliveira e; SCHERRE, Maria Marta Pereira (Orgs.). *Padrões sociolingüísticos: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro*. Rio de Janeiro: Tempo Brasileiro/ Departamento de Lingüística e Filologia – UFRJ, 1996. p. 37-50.

TARALLO, Fernando (1990). *Tempos lingüísticos: itinerário histórico da língua portuguesa*. São Paulo: Ática. p. 57-67, 160-72.

TARALLO, Fernando (1993). Sobre a alegada origem crioula do português brasileiro: mudanças sintáticas aleatórias. In ROBERTS, Ian; KATO, Mary (orgs.). *Português brasileiro: uma viagem diacrônica*. Campinas: Editora da UNICAMP. p. 35-68.

TARALLO, Fernando (1997). *A pesquisa sociolingüística*. 5. ed. São Paulo: Ática.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin I (1968). *Empirical foundations for a theory of language change*. Austin: University of Texas Press.