



“[...] caelle da muytas uezes victoria a aquelles que creem e defende a sua sãcta fé”: concordância verbal no português arcaico ¹

“[...] *caelle da muytas uezes victoria a aquelles que creem e defende a sua sãcta fé*”:
subject-verb agreement in the archaic Portuguese

Pedro Daniel dos Santos Souza*
Universidade do Estado da Bahia,
Eunápolis, Bahia, Brasil

Resumo: Este trabalho discute a variação da concordância verbal na primeira fase do período arcaico da língua portuguesa (séculos XIII e XIV), a partir de *corpus* constituído por textos literários e não-literários (oficiais, particulares e institucionais) representativos da produção medieval portuguesa. Fundamentando-se no aporte teórico-metodológico da Sociolinguística Histórica (ROMAINE, 1982) a análise dos dados confirma 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. 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 e realização do sujeito demonstram desempenhar um papel significativo no uso da regra da concordância verbal. Com base nos resultados, verbo inacusativo, menor nível de saliência, posposição ao verbo, sujeito [- humano], texto literário e sujeito realizado desfavorecem a aplicação da regra de concordância verbal. O comportamento dessas variáveis permite-nos atestar que a variação na primeira fase do período arcaico da língua portuguesa não é aleatória, mas condicionada por fatores que também têm sido considerados relevantes para a variação no português brasileiro (PB).

Palavras-chave: Português arcaico. Concordância verbal. Variação. Regra variável.

Abstract: This paper discusses the variation of subject-verb agreement in the first phase of the archaic period of the Portuguese language (13th and 14th centuries), based on a corpus of literary and non-literary texts (official, private and institutional) representative of medieval Portuguese production. Relying on the theoretical-methodological contribution of Historical Sociolinguistics (ROMAINE, 1982), the analysis of the data confirms the existence of a variable syntactic rule ready in the first phase of the archaic period, defined by their fluence of morphophonetic, syntactic and semantic factors. The verb type variables, phonics alience, position of the subject in relation to the verb, indication of the plural in the subject, semantic characterization of the subject, type of text and realization of the subject demonstrate a significant role in the use of the verbal concordance rule. According to the results, um accusative verb, lower level of protrusion, posposition to the verb, [- human] subject, literary text and realized subjectt end to desfavor the application of the subject-verb agrément rule. Such variables allow us to attest that the variation in the first phase of the archaic period of the Portuguese language is not arbitrary, but rather conditioned by factors that have al so been considered relevant for the variation in Brazilian Portuguese (BP).

Keywords: Archaic Portuguese. Verbal agreement. Variation. Variable rule.

* Universidade do Estado da Bahia (UNEB), Eunápolis, Bahia, Brasil. E-mail: pdsouza@uneb.br

¹ Os dados discutidos no presente artigo foram extraídos de nossa Dissertação de Mestrado, intitulada *Concordância verbal em português: o que nos revela o período arcaico?*, orientada pela saudosa Profa. Dra. Rosa Virgínia Mattos e Silva e defendida em 2005, no Programa de Pós-Graduação em Letras e Linguística da Universidade Federal da Bahia (UFBA).

1 DELIMITANDO O PROBLEMA

Há alguns anos, a variação na concordância verbal² em português tem sido objeto de análises exaustivas, seja em abordagens variacionistas, seja sob um enfoque gerativista. Independentemente do modelo teórico utilizado, as análises até agora realizadas têm priorizado o aspecto sincrônico. No entanto, para uma visão mais completa desse fenômeno linguístico, se faz necessário estudá-lo, exaustivamente, sob outra dimensão – a diacrônica. Com isso, não queremos dizer 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 linguístico em questão.

Visando contribuir com “algumas ideias para a concretização de uma história do português brasileiro”, Mattos e Silva (1998a, p. 40) propõe alguns campos de investigação para entrevermos o seu passado, que, embora no seu processo de realização possam ser desenvolvidos de forma autônoma, estão necessariamente interligados: a) a *reconstrução de uma história social linguística do Brasil*, partindo de fontes históricas múltiplas; b) a *reconstrução de uma sócio-história linguística ou de uma sociolinguística histórica*, que se moverá numa perspectiva de estabelecer correlações entre fatores linguísticos e fatores sociais; c) a *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; e d) a *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. Como discutido pela autora, uma conquista “desse programa de trabalho coletivo” seria a possibilidade de demonstrar se o que tipifica o português brasileiro (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”.

A presente discussão, objetivando atender a proposta explicitada no campo *c*, reflete sobre 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 aqui discutidos 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”.

²A 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 (SN) 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”.

2 CONCORDÂNCIA VERBAL: DA VISÃO “PREÇEITIVA” À VISÃO DESCRITIVA

O embate entre descrição e prescrição há muito que tem marcado a história dos estudos linguísticos. 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 linguística 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 linguístico 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).

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 linguísticos 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. 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 linguístico 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. A par dessas questões, nesta seção trataremos da forma como concordância verbal tem sido abordada na perspectiva dos conceitos de *norma* aludidos.

1.1 A CONCORDÂNCIA VERBAL: O OLHAR “PREÇEITIVO”

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 (MARTINS, 2001). Sobre esses documentos, e consequente 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, reconhecemos 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 inaugurará uma visão prescritiva sobre a língua, na direção de uma norma subjetiva, 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, pautada numa norma subjetiva, pode ser observada logo no prólogo de sua *Grammatica* ao definir “gramatica e suas pârtes” e ressaltar que as abordará “nam segundo convém à ordem da Gramática especulativa, mas como requêre 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 ùa 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 quães fôlgam de aprender. Este nome, moços, é do gênero masculino e do número plurár e da terceira pessoa” (BARROS, [1540] 1971, p. 351).

Como podemos 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, o 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, verificamos 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, observamos 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” (CUNHA; CINTRA, 1985, 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” (CUNHA; CINTRA, 1985, 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 comprovarmos 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 (1985, p. 488) 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”. Em linhas gerais, os gramáticos apresentam os seguintes casos particulares que podem ser considerados como facultativos, embora, em algumas situações, admitam a preferência por uma forma (singular ou 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 quatinho 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 *trainu-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: estamos realmente diante de variação no sentido definido pela sociolinguística 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.

A análise cuidadosa da abordagem feita por Cunha e Cintra (1985) 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. Esse comportamento revela uma visão purista que desconhece a possibilidade de variação, e conseqüentemente de mudança, a que todas as línguas estão sujeitas. Somente um novo olhar, essencialmente descritivo, na direção de uma norma objetiva, possibilitará uma maior compreensão do fenômeno linguístico em discussão, uma vez que a prescrição não revela o verdadeiro comportamento adotado pelos falantes no uso da concordância verbal.

1.2 ESTUDOS SOBRE A CONCORDÂNCIA VERBAL NO PORTUGUÊS ARCAICO

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* (em alemão, *Alt Portugiesisches Elementarbuch*). 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 prazer* (Euf. 358)” (HUBER, [1933] 1986, p. 280).

Observamos então que Huber ([1933] 1986) 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” em suas

³ Quando os exemplos apresentados foram retirados de nosso *corpus* de pesquisa, as letras indicam o texto/documento em que se encontram.

palavras, não é obedecida é que devem merecer o olhar do investigador. Diante disso, passa-se a apresentar essas referidas situações que “chamam a atenção”.

Embora ressalte a regra categórica, Huber ([1933] 1986) faz referência a 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 ideia de coletividade, a posposição do sujeito, entre outros contextos. Como destaca: “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)” (HUBER, [1933] 1986, p. 281).

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 pelo gramático:

- i) E logo... *se ajuntarom* deante Santo Antonio *multidom de pexes grandes e pequenos* (Milagres de Sto. Antonio 1)
- ii) — *toda acristãydade que estã* ã 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 ([1933] 1986) 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 gemte te lamça* de ssy com nojo que de *tyhan* (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, o autor 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 destaca, esses casos, considerados 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 levado em conta, isto é, a inacusatividade ou ergatividade⁴ do verbo em questão. Vale destacar que 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”, observamos, por parte de Huber ([1933] 1986), uma primeira tentativa de explicitar o mecanismo da concordância verbal sem a visão “preçetiva” da tradição gramatical, abrindo portas para uma reflexão metalinguística mais detalhada, coerente, na medida do possível, com uma visão que considera a variação linguística.

Verificamos, 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 linguístico 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. A pesquisadora apresenta uma síntese dos casos de variação documentados no *corpus*, retoma a descrição de Huber ([1933] 1986), contrapondo-a 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;
- 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 Mattos e Silva (1986) 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”

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

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 linguísticas investigadas.

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

(MATTOS E SILVA, 1986, p. 97). A discussão apresentada, por sua vez, 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 sobre a concordância, chegando às seguintes conclusões:

- a) a contiguidade 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). No entanto a análise proposta não se esgota nos aspectos focalizados acima. A pesquisadora passa a examinar outros aspectos, cujos resultados são sumarizados abaixo:

- a) com o SN simples a não contiguidade 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 SN sujeitos compostos de parassinônimos, a concordância é variável, sendo mais frequente o verbo no singular;
- c) com SN sujeitos coordenados não-parassinônimos, é mais frequente o verbo estar no plural;

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).

d) anteposição do verbo aos sujeitos permite a variação.

Observamos, 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 nos 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. Na quarta parte, onde 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, verificamos 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 a pesquisadora deixa 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” (ou seja, de gêneros) dos textos, da tradição medieval portuguesa, como realizamos em Souza (2005).

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 mais uma vez a variação documentada por Huber ([1933] 1986), a pesquisadora infere, a partir do conjunto de dados observados, que a variação na concordância decorria, 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

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.

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 frequente 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 examinada.

Mattos e Silva (1998b)¹¹ rediscute *A concordância verbo-nominal facultativa no português arcaico*, destacando, inicialmente, como a questão se configura no português contemporâneo. Conforme a autora, observamos, 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 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 deprender 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 linguístico. Em seguida, faz referência a Huber ([1933] 1986) que, em sua *Gramática do português antigo*, apresenta informações significativas sobre o problema no período arcaico. Mais uma vez os dados apresentados 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 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

¹¹ 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.

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 linguísticos que favorecem esse fenômeno, a partir da análise de variáveis que envolvem não só fatores sintáticos e semânticos mas também morfo-fônicos, caminho que buscamos perseguir (SOUZA, 2005).

Como podemos depreender dos estudos realizados 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 normatizado¹², ou seja, a partir do século XVI;
- ii) conseqüentemente, apesar de não podermos 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 12).

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 português* (1998)¹³ 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 XVI¹⁴, 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

¹² 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.

¹³ 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.

¹⁴ 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); *Bois de Deleyto* (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 Ioam das Índias* (séculos XIV e XVI); *Crónica Geral de Espanha de 1344* (Ms. do século XV).

português moderno, o traço [humano] a restringe de forma bastante acentuada” (NARO; SCHERRE, 1998, p. 62-63). Além do traço [humano], a análise dos casos de sujeito composto ainda 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;
- 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.

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, argumentam que os fatores controladores da variação no português medieval são os mesmos que controlam a variação no PB, uma vez que “the variable constraints that govern the use of agreement did not change over time; only the input weight changed”¹⁵ (NARO; SCHERRE, 2000, p. 250).

Fundamentando-se na análise do fenômeno variável, Naro e Scherre (2000) 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 europeia, mas também se apoia 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

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

¹⁵ “as restrições variáveis que governam o uso da concordância não mudaram com o passar do tempo, somente mudou o peso do *input*” (tradução nossa).

between adult language speakers of particularly diverse origins and the nativization of Portuguese by communities of such speakers¹⁶ (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.

Observamos, 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 a primeira reflexão metalinguística¹⁷ 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 consequentes 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.

Corroborando com as questões apresentadas nesta seção, destacamos que 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 linguística. O que permite evidenciar a importância do *princípio da uniformidade* da *Teoria da variação e mudança*, sendo o qual “the forces which operated to produce the historical record are the same as those which can be seen operating today”¹⁸ (LABOV, 1975, p. 829).

2 SOBRE A REGRA VARIÁVEL DE CONCORDÂNCIA VERBAL NA PRIMEIRA FASE DO PORTUGUÊS ARCAICO

Como discutimos na seção anterior, a variação na concordância verbal no período arcaico da língua portuguesa já foi objeto de investigações diversas, embora tenhamos atestado a necessidade de ampliar a análise para um *corpus* mais representativo, com vistas a uma descrição exaustiva dos contextos linguísticos que favorecem e/ou desfavorecem esse fenômeno. Nessa direção, na presente seção, apresentaremos nossa pesquisa sobre a

¹⁶ “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” (tradução nossa).

¹⁷ Vale ressaltar a *Grammatica da lingua portuguesa* de Fernão de Oliveira, datada de 1536, como, na verdade, a primeira reflexão metalinguística sobre o português, mas sob uma perspectiva descritiva. Assim, a *Gramática* de João de Barros seria a primeira reflexão metalinguística de cunho prescritivo.

¹⁸ “as forças que operaram no passado para produzir a mudança histórica são as mesmas que operam no presente” (tradução nossa).

variação 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, ou seja, os séculos XIII e XIV. Nossa investigação, no âmbito da Linguística Histórica, fundamenta-se no aporte teórico-metodológico da Sociolinguística Histórica, como propõe Romaine (1982). Na presente seção, faremos uma breve caracterização do *corpus* e, em seguida, uma discussão e análise dos dados, que foram submetidos ao VARBRUL¹⁹.

Cabe-nos destacar que, neste trabalho, descrevemos e analisamos a variação na concordância verbal na primeira fase do período arcaico da língua portuguesa (séculos XIII e XIV), estabelecendo uma relação entre fatores sintáticos, morfo-fônicos e semânticos, ou seja, variáveis linguísticas, e fatores extralinguísticos, como o tempo e o tipo (gênero) de texto. Correlacionamos o fenômeno variável em estudo – a variável dependente *presença/ ausência de marcas formais* na relação entre o SN sujeito e o verbo – às construções variáveis da estrutura linguística, analisando a interferência dos seguintes fatores (variáveis independentes ou explanatórias):

- a) linguísticos: 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) extralinguísticos: tempo, tipo de texto.

A apresentação e discussão do fenômeno linguístico em análise neste trabalho fundamentam-se, como dissemos, na Sociolinguística Histórica (ROMAINE, 1982), uma aplicação do quadro teórico desenvolvido por Weinreich, Labov e Herzog (1968), no artigo *Empirical foundations for a theory of language change*, a estudos da variação linguística em sincronias passadas da língua. Conhecido sob vários rótulos – *Teoria da Variação e Mudança*, *Sociolinguística quantitativa*, *Sociolinguística variacionista*, *Sociolinguística laboviana*, entre outros –, o mencionado quadro teórico desenvolvido por Weinreich, Labov e Herzog (1968)

¹⁹ Como informamos no início, os dados aqui discutidos foram extraídos de nossa Dissertação de Mestrado, orientada pela saudosa Profa. Dra. Rosa Virgínia Mattos e Silva e defendida em 2005, no Programa de Pós-Graduação em Letras e UFBA. Na época, o programa de regra variável VARBRUL era o suporte estatístico utilizado nos estudos sobre a variação linguística. 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 mencionado programa. Após o levantamento dos dados, os mesmos foram codificados para que o programa computacional fizesse o tratamento quantitativo (estatístico e probabilístico) das ocorrências das variantes linguísticas encontradas. O programa VARBRUL, além de fazer o cálculo dos pesos relativos de cada variável independente, apresentava uma seleção estatística dos diversos grupos de variáveis analisados. Inicialmente esta seleção ocorria 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 frequências observadas, era o segundo elemento a influenciar a escolha de uma dada variável. Destacamos ainda que um aspecto importante do programa VARBRUL era que trabalhava com diversos níveis de análises, efetuando-se comparações entre os valores probabilísticos atribuídos aos fatores das variáveis.

assume, como coordenada básica da reflexão sobre a mudança, a *heterogeneidade* normal da língua e, ao mesmo tempo, argumenta contra a ideia de que sistematicidade e variabilidade se excluem. Assim, considera que a variação é inerente a todo processo linguístico, não sendo aleatória, mas comandada por restrições linguísticas e não linguísticas, o que abre caminhos para a possibilidade de sistematizar a variação existente e própria da língua falada, no nosso caso, a língua escrita. O estudo da variação sistemática decorrente dos condicionadores estruturais e sociais discute a concepção de língua enquanto sistema homogêneo, definida por si mesma.

Haja vista que a Linguística Histórica é uma disciplina essencialmente interpretativa (*ars interpretandi*) e cabe ao linguista 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 sociolinguístico limitado. Os materiais linguísticos registrados nesses textos apenas podem constituir a base de um trabalho interpretativo a ser desenvolvido pelo linguista 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, na direção de tentar “hearing the inaudible”²⁰ (LASS, 1997, p. 45).

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

2.1 BREVE CARACTERIZAÇÃO DO CORPUS

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 linguísticos dos manuscritos que editam. 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 de gêneros dos textos (textos literários e não-literários). Considerando as edições e os critérios utilizados, com vistas a garantir dados linguísticos confiáveis, o *corpus* deste trabalho²¹ foi constituído por documentos jurídicos ou notariais (oficiais, particulares e institucionais) e documentos literários em prosa de caráter religioso, apresentados e caracterizados abaixo:

²⁰ “ouvir o inaudível” (tradução nossa).

²¹ 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).

- 1) *Testamento de Afonso II* (TA): documento oficial-régio de inícios do século XIII (1214), 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). 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).
- 2) *Documentos notariais editados por Clarinda de Azevedo Maia* (1986) (DNCM): 22 documentos notariais (documentos de doações, compras, vendas, testamentos, inventários de pessoas falecidas, etc.), distribuídos por duas províncias portuguesas (do Douro Litoral e do Minho) que compõem a região de Entre-Douro-e-Minho. Esses documentos são provenientes dos fundos documentais do Arquivo Nacional da Torre do Tombo (Lisboa) e do Arquivo Distrital de Braga.
- 3) *Documentos notariais editados por Ana Maria Martins* (2001) (DNAM): 70 documentos provenientes de fundos documentais de mosteiros do Noroeste de Portugal e da região de Lisboa. São testemunhos da vida econômica das instituições monásticas e dos que delas dependiam ou se moviam na sua esfera, sendo dos seguintes gêneros: arrendamentos, aforamentos, emprazamentos, vendas, câmbios, partilhas, doações, quitações/renúncias, vedorias, posses, confirmações, procurações, recebimentos, notícias, testamentos, sentenças, demandas, relatos de contendas, avenças/composições/acordos. Esses 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).
- 4) *Foro Real de Afonso X* (FRAX): 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 ação de D. Afonso III e de D. Dinis” (FERREIRA, 1987, p. XV). Trata-se de uma tradução do *Fuero Real*. Neste trabalho, utilizamos a edição crítica de José de Azevedo Ferreira (1987). 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).
- 5) *Orto do Esposo* (ORTO): manuscrito de fins do século XIV. 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). Neste estudo, utilizamos 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.

- 6) *Flos Sanctorum* (FLOS): documento do século XIV que 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. Para o presente trabalho, utilizamos a edição interpretativa apresentada por Américo Venâncio Lopes Machado Filho (2003). Trata-se de “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” (MACHADO FILHO, 2003, p. xxiii), que foi trazido para o Brasil por Serafim da Silva Neto, no ano de 1950, juntamente com a versão mais antiga em português dos *Diálogos de São Gregório* e o *Livro das Aves*.

2.2 DESCRIÇÃO E ANÁLISE DAS VARIÁVEIS QUE CONTROLAM AVARIAÇÃO NO PORTUGUÊS ARCAICO

Nos textos que compuseram o *corpus* de nosso trabalho, encontramos 5.901 ocorrências de P6 que, após codificação, foram submetidas à leitura do VARBRUL. 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. Do total de ocorrências de terceira pessoa do plural, 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 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 (1) a (6):

- (1) – (...) E levou-me pera huũ logar muy deleytoso, qual homem outro nõ poderia {F71vC1} achar segundo meu entendimento. Ca a luz daquel logar e a claridade que hi havia **mostrava** a fremosfera do logar. (FLOS)
- (2) E se leuar *per* força sposa alhea e *ante que* a aya de ueer cũ elha nenhũa cousa for lhy tollecta, todo *quanto* lhy for tollecto cũ elha **aiao** o esposo e a esposa *per* meyo. (FRAX)
- (3) [...] o qual herdãmẽto a my **emtregou** o dito Pedro Heanez e a dita ssa molher; o qual herdãmẽto eu reçẽby *para* o dito moesteyro. (DNCM)
- (4) [...] ca de Jhesu Christo **saac** e **nacem** todalas sabedorias, nõtam solamẽte as sciencias sanctas e diuynaaes mas ajnda as sciencias dos filosofos gentis. (ORTO)
- (5) [...] caelle da muytas uezes victoria a aquelles que **creem** e **defende** a sua sacta fé [...] (ORTO)
- (6) Assy como a chuua e a neue **descende** do ceo e **molha** auõdosamẽte a terra e faz dar frũcto [...] (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 programa computacional 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 eliminados os problemas gerados. 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 programa VARBRUL, que é, justamente, aquela considerada estatisticamente significativa para o entendimento da variação que envolve o fenômeno linguístico em estudo. Embora sejam apresentadas as frequências 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 previamente definidas, o programa descartou duas (as variáveis *língua original do texto* e *tempo*), selecionando as demais na seguinte ordem de relevância: 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. A nossa 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 linguístico em estudo. As hipóteses, a definição das variáveis, a decisão de amalgamar variáveis ou mesmo de realizar cruzamento vão na direção das ações do pesquisador para descrição e análise mais precisas.

2.2.1 Tipo de verbo

A primeira variável selecionada foi *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, subdividimos 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 abaixo permitem identificar algumas diferenças.

Tabela 1 – Variável *tipo de verbo*

| FATOR | FREQUÊNCIA | PESO RELATIVO |
|-------------------------|-------------------|---------------|
| Transitivo | 2.954/3.091 (96%) | .45 |
| Inacusativo ou ergativo | 134/228 (59%) | .09 |
| Inergativo | 1.154/1.190 (97%) | .6 |
| Modal/auxiliar | 505/537 (94%) | .40 |
| Cópula | 836/855 (98%) | .71 |

Fonte: elaborada pelo autor.

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). Podemos 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 observarmos o comportamento da cópula que exhibe 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 da tabela 2, possivelmente, permitem-nos 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 |

Fonte: elaborada pelo autor.

A análise dos dados da tabela acima permite-nos 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, constatamos 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, uma hipótese inicial 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.

2.2.2 Saliência fônica

O fator *saliência fônica*, bastante conhecido na bibliografia sobre o português 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, podemos observar 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 (*anya~anyã, sabba~sabbã*), 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*.

Os dados do português arcaico demonstraram que saliência fônica é uma das variáveis mais significativas, uma vez que foi a segunda a ser selecionada pelo programa. Na tabela 3, podemos observar os resultados referentes a essa variável.

Tabela 3 – Variável *saliência fônica*

| FATOR | FREQUÊNCIA | PESO RELATIVO |
|--------------|-------------------|---------------|
| [+ saliente] | 2.768/2.845 (97%) | .64 |
| [- saliente] | 2.815/3.056 (92%) | .37 |

Fonte: elaborada pelo autor.

Os resultados da tabela acima 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 a desfavorecem, tendo em vista um peso relativo de .37. O mesmo teria sido 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.

2.2.3 Posição do sujeito em relação ao verbo

A terceira variável selecionada pelo programa diz respeito a um aspecto relacionado à ordem dos constituintes nas sentenças: *posição do sujeito em relação ao verbo*. Para essa variável, definimos 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 SN 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 | FREQUÊNCIA | 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 |

Fonte: elaborada pelo autor.

Pelos resultados, verificamos uma influência da coesão estrutural para a compreensão do fenômeno da concordância em português arcaico. A distância e a posposição contribuem para definir esse quadro. Observamos 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, como podemos constatar da análise da tabela abaixo.

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

| FATOR | FREQUÊNCIA | PESO RELATIVO |
|-----------|-------------------|---------------|
| Anteposto | 2.392/2.501 (96%) | .55 |
| Posposto | 643/756 (85%) | .34 |

Fonte: elaborada pelo autor.

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 SN pós-verbais é menor do que com SN pré-verbais (peso relativo de .34). Os dados presentes no *corpus* revelam um uso mais frequente das formas verbais singulares em contextos de sujeitos plurais quando estes estão distantes ou pospostos, confirmando assim a importância da variável *posição relativa*.

2.2.4 Indicação do plural no sujeito

A quarta variável selecionada pelo programa, *indicação do plural no sujeito*, possibilitou identificarmos a atuação das estratégias de pluralização sobre a variável dependente *presença/ausência de marcas explícitas de plural*. Associada à questão da coesão estrutural, a indicação plural no sujeito foi subdividida em três fatores: i) indicação mórfica do plural;

ii) quantificador; e iii) valor semântico do item lexical. Na tabela 6, são apresentados os resultados obtidos a respeito dessa variável.

Tabela 6 – Variável *indicação do plural no sujeito*

| FATOR | FREQUÊNCIA | 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 |

Fonte: elaborada pelo autor.

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 (7) a (9) representam alguns contextos favorecedores da variação no *corpus*, levando em consideração o fator “valor semântico do item lexical”, marcado pelos traços [+ singular, - singularidade]:

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

(8) E quando o todo o poboo o **vyo** fazer seu officio na véspera assi como era custume, **houverom** ende gram prazer. (FLOS)

(9) 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 ta boo pastor e de ta boo senhor. (FLOS)

O contexto apresentado nos exemplos acima também já foi descrito por Huber ([1933] 1986) em sua *Gramática do português antigo*. Embora as ocorrências sejam poucas, são dados relevantes, pois demonstram um uso linguístico marcado pelo valor semântico em detrimento da prescrição formal, estrutural.

2.2.5 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. No presente trabalho, esta variável também se mostrou estatisticamente relevante, tendo em vista que foi a quinta selecionada. No *corpus* aqui analisado, verificamos 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 | FREQUÊNCIA | PESO RELATIVO |
|------------|-------------------|---------------|
| [+ humano] | 4.747/4.942 (96%) | .53 |
| [- humano] | 836/959 (87%) | .35 |

Fonte: elaborada pelo autor.

O traço [+humano] do sujeito exerce influência sobre a aplicação da regra apresentando um peso relativo de .53, como pode ser observado no *corpus*. Verificamos, por seu turno, 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.

2.2.6 Tipo de texto

Trabalhos de pesquisa 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 extralinguística. 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*.

Neste trabalho, utilizamos a variável *tipo de texto*, subdividida em função da diversidade tipológica dos textos que compõem o *corpus*, seguindo a orientação de Romaine (1982), a fim de possibilitar uma análise de variáveis externas, embora manuseando um elemento extremamente linguístico: o texto. Na tabela abaixo, apresentamos os dados referentes a essa variável, sendo a sexta estatisticamente relevante para o programa.

Tabela 8 – Variável *tipo de texto*

| FATOR | FREQUÊNCIA | 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 |
|------|---------------|-----|

Fonte: elaborada pelo autor.

Observamos 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 (1986), 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 (2001) favorecem a aplicação da regra com peso relativo de .72.

Com vistas a melhor compreendermos as relações entre textos literários e não-literários, submetemos os dados ao programa, amalgamando os fatores dessa variável e transformando-os em apenas dois: *textos literários* e *textos não-literários*. Na tabela abaixo, apresentamos os resultados.

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

| FATOR | FREQUÊNCIA | PESO RELATIVO |
|-----------------------|-------------------|---------------|
| Textos não-literários | 2.128/2.216 (96%) | .60 |
| Textos literários | 3.349/3.579 (94%) | .44 |

Fonte: elaborada pelo autor.

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.

Observamos 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 (10) a (14), o que corrobora para a tão baixa frequência de contextos favorecedores da ausência de marcas:

(10) **Sabbham** quãtos este testemoyño **uirẽ** e **léer ouuirẽ** *que* Donna Sancha Esteuaiz, [...] (DNM)

(11) *Testemoyas que presentes foro*: frey Pedro – 18 Steueez, frey Afonso da ordin dos preega-dores; Femam da Veiga e eu Johã – 19 Perez, tabaliõ de Miragaya, [...] (DNAM)

(12) Os *que a esto foro presentes*: Béeyto Periz, alfaiame, Steuã Miguéez, Martin F(.....)dj, al – 17 fayates de Bragáá, e Johã Dominguz, clerigos do dito juyz. (DNAM)

(13) **Conhoscam** *quantos este prazo uirẽ e léér ouuirẽ* que eu Stevam anes Abbade do monsteiro deCety / 2 e o Priol e cõuento desse logar. ffazemos [...] (DNAM)

(14) **Sabhã** todos *que em pressença de mj ffrancisco gíral dez publico* tabaliõ de Gujmarães e das *testemunhas que Adeante som* scritas e na crasta / 2 de *santa Maña* [...] (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.

2.2.7 Realização do sujeito

A última variável estatisticamente relevante relaciona-se com o parâmetro *pro-drop*, ou seja, o parâmetro do sujeito nulo. Em termos gerais, 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 abaixo:

Tabela 10 – Variável *realização do sujeito*

| FATOR | FREQUÊNCIA | PESO RELATIVO |
|-------------------|-------------------|---------------|
| Sujeito realizado | 3.031/3.253 (93%) | .22 |
| Sujeito nulo | 2.550/2.646 (96%) | .82 |

Fonte: elaborada pelo autor.

É interessante observarmos que, embora a porcentagem entre os fatores seja aproximada – 93% de ocorrências de sujeito realizado e 96% de sujeito nulo –, o comportamento dos fatores, levando em conta 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 de nossa pesquisa. A hipótese de que “marcas levam a marcas” e “zeros a zeros” não pode ser confirmada à luz dos dados utilizados no presente estudo, 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 linguístico estudado.

3 CONSIDERAÇÕES FINAIS

A variação na concordância verbal em português, um dos fenômenos linguísticos 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 linguísticas 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?* 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 *Grammatica da língua portuguesa*, 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, abrimos novas perspectivas para o entendimento mais amplo da variação que envolve a concordância verbal de número na língua portuguesa.

Este trabalho também permite-nos reafirmar que a variação na concordância verbal não é um fenômeno restrito ao português contemporâneo. 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-nos verificar a relevância de sete 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 e realização do sujeito.

Constatamos 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 verbal. Os dados analisados revelaram que 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 brasileiro, demonstrando ser possível que “as forças que atuam no presente são as mesmas que atuaram no passado” (LABOV, 1975, p. 829).

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. Assim, a *posição relativa* parece ser significativa para compreendermos a variação.

Outro grupo de fatores que desempenha um papel significativo é a *caracterização semântica do sujeito*. O traço [+ humano], da mesma forma que no PB contemporâneo, 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.

Finalmente, 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.

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 extralinguísticos, como a natureza dos textos. Esta investigação representa apenas um passo para a compreensão desse fenômeno linguístico 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 brasileiro 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 os 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. *Grammatica da lingua portuguesa*. Edição de M. L. Buescu. Lisboa: Faculdade de Letras, [1540] 1971.

COSTA, Pe. Avelino de Jesus da. 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, 1979. p. 307-321.

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

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

HUBER, Joseph. *Gramática do português antigo*. Lisboa: Gulbenkian, [1933] 1986.

LABOV, William. On the use of the present to explain the past. *Linguistics at the crossroads*. Liviana: EditriceJupiter Press, 1975.

LASS, Roger. *HistoricalLinguisticsandLanguageChange*. Cambridge: Cambridge University Press, 1997.

LEMLE, Mirian; NARO, Anthony Julius. *Competências básicas do português*. Rio de Janeiro: Fundação MOBREAL; Fundação Ford, 1977.

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

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

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

MAIA, Clarinda de Azevedo. *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, 1986.

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

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

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

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

MATTOS E SILVA, Rosa Virgínia. *Estruturas trecentistas: elementos para uma gramática do português arcaico*. Lisboa: IN-CM, 1989.

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

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

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

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

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

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

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

ROMAINE, Suzzane. *Socio-historical linguistics: its status and methodology*. Cambridge: Cambridge University Press, 1982.

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

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

SOUZA, Pedro Daniel dos Santos. *Concordância verbal em português: o que nos revela o período arcaico?* 2005. Dissertação (Mestrado em Letras e Linguística) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2005.

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

Recebido em: 20/12/2020

Aprovado em: 05/02/2021

Publicado em: 25/07/2021